DOI: 10.59277/PL.2025.28.13

A NOTE ON THE UNSAYABLE: USING SPLIT LANGUAGES FOR EPISTEMIC INCOMPLETE DESCRIPTIONS

MARIAN CĂLBOREAN

Faculty of Philosophy, University of Bucharest

Abstract. This note introduces the core intuition and basic construction of split epistemic languages, a novel framework for modeling incomplete descriptions in epistemic logic. Standard Kripke models often overstate what agents can plausibly formulate by presupposing a shared maximal language. The proposed remedy is to assign each agent a sub-language called "split language", generated from only those concepts which they can grasp. By blocking inexpressible formulas at the syntactic level, this approach distinguishes lack of information (uncertainty) and lack of conception (unawareness), solving some standard puzzles of epistemic logic. This note presents the initial problem statement, the core intuition and basic formal machinery, together with several applications which demonstrate the framework's applicability.

Keywords: epistemic logic, awareness, expressibility, uncertainty, muddy children puzzle.

1. INTRODUCTION

It is well-recognized that standard epistemic logic assumes a single common model, introduced at the moment of formalization. We observe that there are really two tacit assumptions doing work: the well-known common knowledge of the model's structure, but also, more subtly, the fact that the logic assumes common use of the standard language, also shared by all agents. But agents often have no fact of the matter about certain concepts, so the most elegant way to model such unawareness is to tie each agent to a language containing only those conceptual tools they have a fact of the matter about. As Ludwig Wittgenstein famously wrote at the end of the *Tractatus*: "Whereof one cannot speak, thereof one must be silent".

If we adopt this novel way of distinguishing lack of information (uncertainty) from lack of conception (unawareness), we have a syntactic way to solve standard puzzles of epistemic logic, by simply identifying the sub-language the agent had been using.

The next sections reproduce almost verbatim the initial conceptual sketch for the split-languages framework, first circulated by the author in 2023 during a scholarship at the CUNY Graduate Center. We first outline the problem, then present the core intuition, followed by the basic formal construction and several simple applications to canonical puzzles[1].

---

[1] For a full, rigorous formalization of this framework, including a proof system, meta-theoretic results, and dynamic extensions, see the author's forthcoming paper, 'What We Cannot Say: An Agent-Relative, Syntactic Account of Ignorance in Epistemic Logic," soon available on arXiv.

2. PROBLEM AND CORE INTUITION

As Sergei Artemov writes, describing the limits of standard epistemic logic:

> "Given a verbal description of a situation, a typical epistemic user cherry-picks a 'natural model' [...] and regards it as a formalization of the original description.". [...] Fundamental deficiencies:
> 1. It covers only complete descriptions, whereas many (intuitively most) epistemic situations are partially described and cannot be adequately specified by a single model.
> 2. The traditional epistemic reading of Kripke/Aumann models tacitly requires common knowledge of the model" [2]

More specifically, also by Artemov:

> "A set of formulas Γ is complete if for each formula F, $\Gamma \vdash F$ or $\Gamma \nvdash \neg F$.
> Examples (two agents with modalities K_1 and K_2:
> 1. *Ignorance*: $\Gamma = \{m\}$, where m is a propositional letter not known to agents. Neither $K_i m$ nor $\neg K_i m$ is derivable, Γ is incomplete.
> 2. *Mutual Knowledge*: $\Gamma = \{K_1 m, K_2 m\}$, both agents have first-order knowledge of m. However, second-order knowledge assertions, e.g., $K_2 K_1 m$, are independent. Γ is incomplete."[3]

First of all, two definitional observations. "Complete description" under supervaluationism means semantical definability which comes down to specifying semantic values for all the parameters (propositional letters, epistemic operators). That is, the model will contain all the possible states of all. But also, "common knowledge" under supervaluationism means that that there is no way to specify the viewpoint of a specific epistemic agent, since all the agents use the complete description given above.

Note that the problem is not that there is a "view from above", namely that of the observer, the problem is that we do not have other expressible views.

Thus, the minimal requirement to solve the Artemov's puzzle is two-fold:

a) Model the idea that there is no fact of the matter as to whether some parameters are true or false for an agent.
b) Express the viewpoints of specific agents, which are partial as compared to the all-knowing observer.

The simplest solution is to *tie a language to an agent: the language of an agent only contains the parameters of which there is a fact of the matter* (not propositions whose letter is not known to him/her, neither other agents which are not known to him/her). And this idea is quite intuitive, in a Wittgensteinian way: if I don't know of the existence of a proposition, I cannot speak of it, I should not be in my language.

This intuition leads directly into a formal construction.

[2] S. Artemov Lecture 12, Nov 28 2023 p2, LFCS, CUNY GC.

[3] S. Artemov Lecture 12, Nov 28 2023 p7, LFCS, CUNY GC.

3. BASIC CONSTRUCTION

3.1. The *epistemic language* L will be built from:

a) V_L – Vocabulary of L, = $<L_p, L_l, L_m>$:
L_p – Propositional operators: $\{\neg, \wedge, \ldots\}$
L_l – Propositional letters: $\{a, b, c, \ldots\}$
L_m – Modal operators, one per agent: $\{K_1, K_2, \ldots\}$

b) r – Well-formation rules (standard)

So, we will write L = r(V). L is the set of all formulas obtainable from the vocabulary by the rules.

3.2. M is a *split epistemic language ("sel")* of L iff it uses the same rules and propositional operators as L and its propositional letters and modal operators are included in those of L:

$V_M = <M_p, M_l, M_m>$ and $M_p = L_p$, $M_l \subseteq L_l$, $M_m \subseteq L_m$
$M = r(V_M)$

Then $M \subseteq L$ and L itself is a sel of L.

3.3. A *split epistemic Kripke model* ("seKm") is a fourtuple $M = <W, <R_1 \ldots R_n>, \models, S>$ of L with:

a) $<W, <R_1 \ldots R_n>, \models>$ form a standard epistemic Kripke model
b) S is a tuple of split epistemic languages of L.

3.4. We define *derivation according to a language*, we allow it to be incomplete:

$\Gamma \vdash_G \varphi$, read "φ is derivable from Γ according to language G", if $\Gamma \vdash \varphi$ and $\Gamma \cup \{\varphi\} \subseteq G$
$\Gamma \nvdash_G \varphi$, read "φ is not derivable from Γ according to language G", if $\Gamma \nvdash \varphi$ and $\Gamma \cup \{\varphi\} \subseteq G$

Note that for L, the language of a seKm, $\vdash_L$ iff $\vdash$.

3.5. Define *consequence according to a language*:

$\Gamma \models_G \varphi$, read "φ is a consequence of Γ according to language G", iff $\Gamma \models \varphi$ and $\Gamma \cup \{\varphi\} \subseteq G$

Note that for L, the language of a seKm, $\models$ iff $\models_L$.

4. APPLICATIONS

4.1. Ignorance and partial knowledge

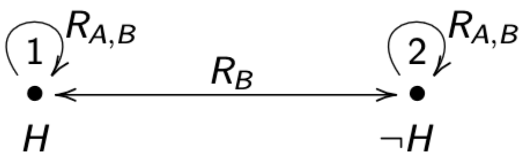
a) Suppose: "Ignorance = $\Gamma_1 = \{p\}$"
That is, p is true but agents know nothing about p.

Then, the languages of the agents should not contain p.

b) Suppose: "Mutual Knowledge = $\Gamma_2 = \{K_1p, K_2p\}$"
That is, no agent knows about what the other knows.

Then, the language of an agent should not contain the epistemic operator corresponding to the other agent.

4.2. The coin – simple case
Suppose: "A coin is tossed and lands heads up. Alice sees the coin, Bob does not"



The seKm will be along the lines of $\langle W, \langle R_A, R_B\rangle, \vDash, S)$ of L, with S containing:

S_A = the language of Alice, $r(\langle L_p, \{H\}, \{K_A\})$ - Alice has no fact of the matter of what Bob knows.

S_B = the language of Bob, $r(\langle L_p, \{H\}, \{K_B\})$ - Bob has no fact of the matter of what Alice knows.

Note that $L = r(\langle L_p, \{H\}, \{K_A, K_B\})$.

It should be immediate that:

a) For the observer it is true that Alice knows that Bob does not know H and that Bob knows that Alice knows whether H
$\vDash K_A \neg K_B H$
$\vDash K_B K_A^W H$[4]

b) But not according to the language of Alice or Bob, since the formula is outside those languages
$\nvDash_{SA} K_A \neg K_B H$
$\nvDash_{SA} K_B K_A^W H$
$\nvDash_{SB} K_A \neg K_B H$
$\nvDash_{SB} K_B K_A^W H$

c) On the derivation side, for Alice:
\`$\vdash_{SA} K_A \neg K_B H$\` does not hold, since \`$K_A \neg K_B H$\` $\notin S_A$
\`$\nvdash_{SA} K_B K_A^W H$\` does not hold, since \`$K_B K_A^W H$\` $\notin S_A$

And the same for Bob.

---

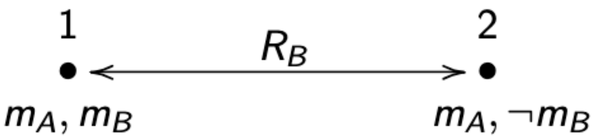
[4] \`$K_A^W \varphi$\` is short for \`$K_A \varphi \vee K_A \neg \varphi$\`.

4.3. Muddy Children Explicit

"A group of n children meet their father after playing in the mud. Each child sees everybody else's foreheads. The father says: "k of you are muddy" after which it becomes common

knowledge that each child knows whether (s)he is muddy. Why?"

Let's take the case of n=2, Alice and Bob, $MCE_{2,2}$ with the assumption 'Alice knows that she is muddy'.



The seKm is along the lines of <W, <R_A, R_B>, ⊨, S) of L.

L is r(<L_p,{m_A, m_B},{K_A,K_B}) where m_A is 'Alice is muddy' and m_B is 'Bob is muddy'

S contains:

S_A = the language of Alice knowing whether she is muddy but not knowing about what Bob knows, r(<L_p,{m_A,m_B},{K_A})

S_B = the language of Bob not knowing whether he is muddy and not knowing about what Alice knows, r(<L_p,{m_A},{K_B})

It should be immediate that:

a) For the observer it is true that Alice knows that Bob does not know that he is muddy

$\vDash K_A \neg K_B m_B$

b) But not according to the language of Alice or Bob, since the formula is outside those languages

$\nvDash_{SA} K_A \neg K_B m_B$

$\nvDash_{SB} K_A \neg K_B m_B$

c) Alice cannot derive neither that Bob knows that he is muddy nor that he does not

\`$\vdash_{SA} K_B m_B$\` does not hold, since \`$K_B m_B$\` $\notin S_A$

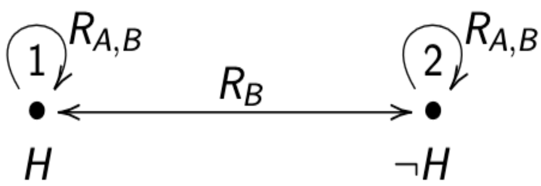
\`$\nvdash_{SA} K_B m_B$\` does not hold, since \`$K_B m_B$\` $\notin S_A$

But once the father makes his announcement, both Alice and Bob switch to using L.

5. COMPLEX CASES – THE PROBLEM OF UNDERSPECIFICATION

Artemov also presents the problem of underspecification[4], i.e. in the coin-tossing example we intend:

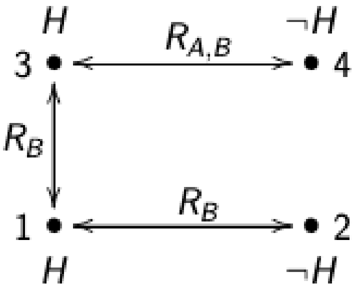
(i) "It is possible for B that A knows ¬H", $\vDash \neg K_B \neg K_A \neg H$



But without

(ii) "Bob knows that Alice knows whether H", $\nvDash K_B K_A^w H$

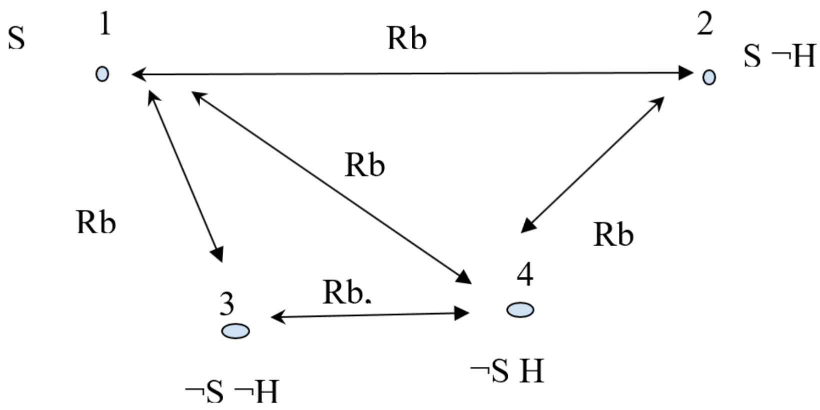
Which has this counterexample:



To model this with split languages, the intuition is: there is a coin toss and, from Bob's perspective, the coin may have been seen by Alice or not. If the coin has been seen by Alice, there should be a fact of the matter for Bob whether Alice knows something, otherwise no.

With S = *the coin was seen by Alice*, the model becomes:

[4] S. Artemov Lecture 12, Nov 28 2023 p40, LFCS, CUNY GC



R_a is reflexive and holds between 3 and 4 (if the coin was not seen, for Alice both H and ¬H are possible). R_b is total.

We have two languages for Bob:

Worlds 1, 2:
$S_{B+} = r(\langle L_p,\{S,H\},\{K_A,K_B\})$ - Bob has fact of the matter of what Alice knows. This is L in fact.

Worlds 3, 4:
$S_{B-} = r(\langle L_p,\{S,H\},\{K_A\})$ - Bob has no fact of the matter of what Alice knows.

Now:
(i) "It is possible for B that A knows ¬H", $\vDash_{SB+} \neg K_B \neg K_A \neg H$ follows at 1 (in the language of Bob with a fact of the matter on A's knowledge)
(ii) "Bob knows that Alice knows whether H", $\nvDash_{SB+} K_B K_A^w H$ does not follow at 1 (in the language of Bob with a fact of the matter on A's knowledge) because $1R_b4$ where $4 \nvDash K_A^w H$

6. CONCLUSION

The framework of split languages, i.e. sub-languages of the standard language of epistemic logic, offers a promising direction for modeling unawareness. By treating an agent's expressive capacity as a primitive, limited only to the concepts available to them, we distinguish what an agent is uncertain about from what they cannot even conceive, with a simple formal implementation, since the distinction is strictly syntactic. And we showed how this approach solves transparently several classic puzzles without presupposing common knowledge of a complete model.

Therefore, the fundamental shift proposed here is to relocate the problem of unawareness from complex semantic machinery to a simple, generative syntax based on an agent's available concepts. Open issues for this novel approach include metalogical results, the proof system, and dynamic extensions for modeling concept acquisition, which will be presented in a larger, forthcoming work.

BIBLIOGRAPHY

1. Artemov, S. (2023). *Lecture* 12, LFCS, CUNY GC. November 28, 2023.
2. Fagin, R., Halpern, J.Y., Moses, Y., & Vardi, M.Y. (1995). *Reasoning About Knowledge*. MIT Press.
3. Fagin, R., & Halpern, J.Y. (1987). *Belief, Awareness, and Limited Reasoning.* „Artificial Intelligence", 34(1), 39–76.
4. Wittgenstein, L. (1992). *Tractatus Logico-Philosophicus*. Kegan Paul, Trench, Trubner.