

# Logic of Non-monotonic Interactive Proofs<sup>\*</sup>

Simon Kramer

University of Luxembourg  
simon.kramer@a3.epfl.ch

**Abstract.** We propose a monotonic logic of internalised *non-monotonic* or *instant* interactive proofs (LiiP) and reconstruct an existing monotonic logic of internalised monotonic or persistent interactive proofs (LiP) as a minimal conservative extension of LiiP. Instant interactive proofs effect a *fragile* epistemic impact in their intended communities of peer reviewers that consists in the *impermanent* induction of the knowledge of their proof goal by means of the knowledge of the proof with the interpreting reviewer: If my peer reviewer knew my proof then she would *at least then* know that its proof goal is true. Their impact is fragile and their induction of knowledge impermanent in the sense of being the case possibly only at the instant of learning the proof. This accounts for the important possibility of internalising proofs of statements whose truth value can vary, which, as opposed to invariant statements, cannot have persistent proofs. So instant interactive proofs effect a *temporary* transfer of certain propositional knowledge (knowable *ephemeral* facts) via the transmission of certain individual knowledge (knowable *non-monotonic* proofs) in distributed systems of multiple interacting agents.

**Keywords:** agents as proof- and signature-checkers, constructive Kripke-semantics, interpreted communication, multi-agent distributed systems, interactive and oracle computation, proofs as sufficient evidence.

## 1 Introduction

The subject matter of this paper is modal logic of interactive proofs, i.e., a novel logic of *non-monotonic* or *instant* interactive proofs (LiiP) [1] as well as an existing logic of monotonic or persistent interactive proofs (LiP) [2]. (We abbreviate interactivity-related adjectives with lower-case letters.) The goal here is to define LiiP axiomatically and semantically as well as to reconstruct LiP as a minimal conservative extension of LiiP. So for distributed and multi-agent systems, whose states and thus truth of statements about states can vary, proof non-monotonicity (as in LiiP) is in a logical sense more primitive than proof

---

<sup>\*</sup> Work funded with Grant AFR 894328 from the National Research Fund Luxembourg cofunded under the Marie-Curie Actions of the European Commission (FP7-COFUND), and finalised during an invited stay at the Institute of Mathematical Sciences, Chennai, India.

monotonicity (as in LiP). In contrast, proof monotonicity is perhaps more intuitive than proof non-monotonicity within formal physical theories validated by experiment and surely within mathematical theories known to be consistent.

Rephrasing [3, Section 1.1] model-theoretically, the proof modality of LiiP internalises a non-monotonic notion of proof in the sense that it can happen that a proposition  $\phi$  can be proved with a (non-monotonic) proof  $M$  to an agent  $a$  in some system state  $s$ , but not anymore in some subsequent state  $s'$  in which  $a$  will have learnt additional or lost previously learnt data  $M'$ . See [1] for formal application examples. Like in LiP [2], we understand interactive *proofs as sufficient evidence* to intended *resource-unbounded* (though unable to guess) proof- and signature-checking agents (designated verifiers).

*Instant* interactive proofs effect a *fragile* epistemic impact in their intended communities  $\mathcal{C}$  of peer reviewers that consists in the *impermanent* induction of the (propositional) knowledge (not only belief) of their proof goal  $\phi$  by means of the (individual) knowledge of the proof (the sufficient evidence)  $M$  with the designated interpreting reviewer  $a$ : If  $a$  knew my proof  $M$  of  $\phi$  then she would *at least then* know that the proof goal  $\phi$  is true. By individual knowledge we mean knowledge in the sense of the transitive use of the verb “to know,” here to know a message, such as the plaintext of an encrypted message. Notation:  $a \mathbf{k} M$  for “agent  $a$  knows message  $M$ ” (cf. Definition 1). This is the classic concept of knowledge *de re* (“of a thing”) made explicit for messages, meaning taking them apart (analysing) and putting them together (synthesising). Whereas by propositional knowledge we mean knowledge in the sense of the use of the verb “to know” with a clause, here to know that a statement is true, such as that the plaintext of an encrypted message is (individually) unknown to potential adversaries. Notation:  $K_a(\phi)$  for “agent  $a$  knows that  $\phi$  (is true)” (cf. Fact 1). This is the classic concept of knowledge *de dicto* (“of a fact”).<sup>1</sup> (We distinguish individual and propositional knowledge with respect to the “*object*” of knowledge [the known], i.e., with respect to a message and clause, respectively. However, individual as well as propositional knowledge can both be individual with respect to the *subject* of knowledge [the knower], i.e., an [individual] agent.) With respect to belief, propositional knowledge essentially differs in that it is necessarily true whereas belief is possibly false, as commonly known and accepted [4]. The epistemic impact of our instant interactive proofs is fragile and their induction of knowledge impermanent in the sense of being the case possibly only at the instant of learning the proof. This accounts for the important possibility of internalising proofs of statements, whose truth value can vary, such as statements about system states, which, as opposed to invariant statements, cannot have persistent proofs. Proofs must (not) prove true (false) statements! Standard examples of statements of variable truth value are contingent (e.g., elementary) facts (expressed as atomic formulas) and characteristic formulas of states [5].

In contrast [2], the epistemic impact of *persistent* interactive proofs is *durable* in the sense of being the case necessarily at the instant of learning the proof *and henceforth*, where time can be present implicitly (such as here) or explicitly

<sup>1</sup> In a first-order setting, knowledge *de re* and *de dicto* can be related in Barcan-laws.