Inaugural Steps in a Computational Study of Time Travel*

⋆ extended abstract ⋆

Atriya Sen • Naveen Sundar G. • Selmer Bringsjord
Rensselaer Artificial Intelligence and Reasoning (RAIR) Lab
Department of Computer Science; Department of Cognitive Science
Rensselaer Polytechnic Institute (RPI)
Troy NY 12180 USA

26 February 2017

We explain inaugural steps in a new, formal, computational study of the possibility of 'time travel,' the ultimate goal of which is to conclusively settle, by machine-verified proof, whether or not human time travel to the past is possible.

An inaugural step is to formalize the concepts of causality and causal loops in a novel *event calculus*: the *Relativistic Event Calculus* (REC). Event calculi are a family of first-order calculi used to model common-sense reasoning in AI [Mue06]. REC, in the spirit of Pat Hayes's seminal "naïve-physics" programme [Hay78, Hay85], constrains logical possibility with physical axioms[1].

With REC and other formal tools in hand, our study of time travel targets paradoxes of causality, including specifically the well-known "looping painter" paradox [Dum86], or what we dub the Paradox of Proust (PoP), which describes a situation where a painting is transported back in time, and is merely copied by its supposed originator. This is an information/creativity paradox, where the existence of the painting is apparently inexplicable. Our argument for possibility is modal, and moreover here both logical and physical possibility must be explicitly considered. Therefore, we augment the REC with a modal system, resulting in a quantified modal logic. We then employ a novel automated theorem prover, ShadowProver[2], to formalize and study, by means of automated and semi-automated reasoning, arguments proposed in the literature for the resolution of this paradox.[3]

*This material is based upon work supported by the Air Force Office of Scientific Research under award number FA9550-17-1-0191.

[1]For now, by *logical possibility*, we mean in this context merely the absence of logical contradictions (clearly, in *all* possible worlds). This concept is to be distinguished from the modal notion of metaphysical possibility, denoting validity in *some* possible world. Physical possibility, at least for now, denotes satisfiability when combined with physical axioms.

[2]See [BG17] for an application.

[3]For example, Horwich's argument [Hor87] for the *improbability* of (backward) time travel,

A Brief Introduction to the Event Calculus and our *Relativization*

The event calculus[a] is a theory expressed in multi-sorted first-order logic. The theory has the two fundamental sorts: Event and Fluent. The Event sort is used to talk about events and actions; the Fluent sort is used to talk about states of the world. Included are the following predicate symbols:[b]

1. *Happens* : Event × Time → Boolean, for declaring when an event happens;
2. *HoldsAt* : Fluent × Time → Boolean, for asserting when a fluent holds;
3. *Initiates* : Event × Fluent × Time → Boolean, for declaring when an event initiates a fluent; and
4. *Clipped* : Time × Fluent × Time → Boolean, for declaring that a fluent becomes not-true between two time points.

One fundamental axiom states that when an event that initiates a fluent happens and nothing clips that fluent, the fluent holds. As can be seen, this axiom, like all axioms in the event calculus, is abstract and has precious little physics content in it:

$$\forall e : \mathsf{Event}, f : \mathsf{Fluent}, t_1, t_2 : \mathsf{Time}. \begin{bmatrix} Happens\,(e, t_1) \wedge \\ Initiates\,(e, f, t_1) \wedge \\ t_1 < t_2 \wedge \\ \neg Clipped\,(t_1, f, t_2) \end{bmatrix} \Rightarrow HoldsAt\,(f, t_2)$$

The *relativistic* event calculus fixes this issue by modifying the event calculus axioms by adding a bridge to physics in the form of conditions Φ that constrain the events and fluents:

$$\forall e : \mathsf{Event}, f : \mathsf{Fluent}, t_1, t_2 : \mathsf{Time}. \left(\Phi\Big(e, f, t_1, t_2\Big) \wedge \begin{bmatrix} Happens\,(e, t_1) \wedge \\ Initiates\,(e, f, t_1) \wedge \\ t_1 < t_2 \wedge \\ \neg Clipped\,(t_1, f, t_2) \end{bmatrix} \right) \Rightarrow HoldsAt\,(f, t_2)$$

With Φ suitably deduced from the axioms of a relativistic theory, it becomes possible, for example, to unify the concept of an event, central to event calculi, with its relativistic interpretation.

[a] We use the singular here, because we are talking about one commonly used specimen in a large family.

[b] The full event calculus has many more symbols.

bilking arguments against reverse-causation, quantum-mechanical arguments postulating multiple worlds, questions of personal identity, and Gödel's argument for the ideality of time [Fef89].

References

[BG17] Selmer Bringsjord and Naveen Sundar Govindarajulu. Are Autonomous-and-Creative Machines Intrinsically Untrustworthy? In Hussein Abbass, Jason Scholz, and Darryn Reid, editors, *Foundations of Trusted Autonomy*. Springer, Berlin, Germany, 2017.

[Dum86] Michael Dummett. Causal Loops. In Raymond Flood and Michael Lockwood, editors, *The Nature of Time*, pages 135–169. Oxford University Press, Oxford, UK, 1986.

[Fef89] S Feferman. A remark about the relationship between relativity theory and idealistic philosophy. *Kurt Gödel: Collected Works*, 2, 1989.

[Hay78] Pat Hayes. The naïve physics manifesto. In D. Ritchie, editor, *Expert Systems in the Microeletronics Age*, pages 242–270. Edinburgh University Press, Edinburgh, Scotland, 1978.

[Hay85] Pat J. Hayes. The Second Naïve Physics Manifesto. In J. R. Hobbs and B. Moore, editors, *Formal Theories of the Commonsense World*, pages 1–36. Ablex, Norwood, NJ, 1985.

[Hor87] Paul Horwich. *Asymmetries in Time*. MIT Press, Cambridge, MA, 1987.

[Mue06] E. Mueller. *Commonsense Reasoning: An Event Calculus Based Approach*. Morgan Kaufmann, San Francisco, CA, 2006. This is the first edition of the book. The second edition was published in 2014.