Logic meets Wigner's Friend (and their Friends)

Alexandru Baltag Sonja Smets
ILLC, University of Amsterdam, the Netherlands
thealexandrubaltag@gmail.com S.J.S.Smets@uva.nl

The abstract for this talk at QPL 2026 is based on [6]: https://arxiv.org/abs/2307.01713
In this presentation we look at the interplay of multi-agent epistemic logic and quantum mechanics. We show under which conditions a logically consistent framework can be provided that satisfies both the standard rules and axioms of multi-agent epistemic reasoning and the axiomatic theory that regulates the flow of quantum information. We take a fresh look at Wigner's Friend thought-experiment [31] and some of its more recent variants and extensions [15, 10] , such as the Frauchiger-Renner (FR) Paradox [18]. The questions we answer include: What is the correct epistemic interpretation of the multiplicity of state assignments in these scenarios? Under which conditions can one include classical observers into the quantum state descriptions, in a way that is still compatible with traditional Quantum Mechanics? Under which conditions can one system be admitted as an additional 'observer' from the perspective of another background observer? By answering these questions we provide a principled solution to Wigner Friend-type paradoxes.

1 Introduction

Wigner's Friend thought-experiment [31], and its variations [15, 10] and extensions [18] seem to indicate that, if quantum theory is assumed to be universally valid (and hence can be applied to complex systems composed of quantum systems as well as 'classical observers'), then different agents are rationally entitled to ascribe different (pure) states to the same system, and as a result they cannot share their information in a consistent manner. More precisely, the result in [18] is stated in the format of a no-go theorem, stating that any theory which satisfies conditions (Q) (the universal validity of quantum theory), (C) (the consistency between agents about their predictions of measurement outcomes), (S) (the unique outcomes of measurements), will lead to a contradiction.

The community replied to this problem with over 100 papers, some of which do refer to the use of modal logic and epistemic reasoning (see e.g. [26, 8, 14]). However, none of these logic-based approaches actually shows what an adequate constraint on the merged (modal-epistemic and quantum) axiom systems should be to obtain a consistent set of axioms to reason about all assumptions (S,C,Q) together. In contrast, our work in [6] provide a clear-cut criterion for *when is epistemic reasoning applicable to 'agents' that are quantum-physical systems.*

The setting we work with, assumes the standard Hilbert-space formalism of quantum theory (as formalized in [34]), as well as the common-sense use of higher-order level reasoning of classical agents (as formalized in basic Epistemic Logic [19]). Further, we follow both Heisenberg [33] and von Neumann [34] by specifying a clear dividing line, a *cut* c_{SO} between what counts as the system S under observation and what counts as the 'observer' O. But when multiple observers are at play and systems S themselves can contain observers, we have to specify where to place the cut for all observers. Moreover, for a given observer O with a specific cut c_{SO}, when is O entitled to admit some (other part of the) observed system S as an *additional observer* O' (who comes with her own cut $c'_{O'S'}$), in such a way that it is still consistent with the predictions made by standard quantum theory? Our answer is given by extending the relational view of Quantum Mechanics to the property of "being an observer", which becomes itself relativized. More specifically, we define a relational notion of *admissible observer* that is always relative both to another background observer O and to a given history h or protocol π. Where a history h is understood to be a sequence of linear maps performed on the state of a given system during the run of an experiment and a protocol π is taken to be a family of histories (both notions are defined below). Roughly speaking, a system A is then called an "admissible observer" with respect to O and h if and only if it is always possible that (as far as the background observer O can know) that none of the information carried by A will be fully erased from the universe at any moment of the given history h.[1] This can happen either because A's memory stays intact until the end of h, or because the information keeps "leaking" out of A, e.g. in the form of copies or records that are being disseminated and copied again before they can be destroyed. A further step we require is the *mutuality* of this relationship: a family of subsystems forms a "community of admissible observers" if each of them is an admissible observer with respect to all the others. What we argue for in [6] is that full higher-order epistemic reasoning (of observers reasoning about other observers etc.) is consistently applicable only within such communities of admissible observers.

We show in [6] that typical Wigner's Friend-type scenarios break our conditions: some of the supposed "agents" (such as the Friend inside the Lab) are *not* admissible observers from the perspective of the other agents (such as Wigner outside the Lab), precisely because the later ones get to *know* that the information carried by the earlier one is being completely erased from the universe. In contrast, in the epistemic scenarios that underly standard multi-agent protocols in Quantum Information and Quantum Computation theory, one can claim that the conditions required for mutual admissibility as observers *are* satisfied.

2 Formalization

We are given a (finite or infinite) list of labels (letters) $\Sigma = \{s_1, s_2, \ldots\}$, denoting *elementary systems*, together with associated Hilbert spaces $H^{(s)}, \ldots$. From these, we can generate composite *systems*, given by subsets $S, O, S', A, B, \ldots \subseteq \Sigma$, each with its associated Hilbert space $H^{(S)} := \bigotimes_{s \in S} H^{(s)}$. We leave unspecified *which* subsets $S \subseteq \Sigma$ are admitted as "systems", in order to accommodate various approaches: according to one approach *all* subsets are systems, while others may impose some restrictions (e.g. finiteness).

The *state* of any system $S \subseteq \Sigma$ from the perspective of some 'observing' system $O \subseteq \Sigma$ is denoted by S_O. When we want to make explicit the time of this state attribution, we write S^t_O (for the state of S at time t from the perspective of O), where we assume given a discrete set of temporal moments $t = 0, 1, \ldots$. All systems S (including Σ) are always described from such a 'situated' point of view, associated to a background observer O. The described system S may or may not overlap with the 'observing' system O, or we may even have $O \subseteq S$. In case they do overlap, the "observed subsystem (of S)" is the non-overlapping part $S - O$ of the system under description.

We allow the 'observer' O to describe systems S that it may not observe (since they may include parts $S \cap O$ that overlap with itself). While it is commonly assumed that no system can observe itself, there is no reason to deny an observer the capacity to 'know' or represent its own state. Indeed, if we identify this state with the observer's *memory* (e.g. the record of all O's measurement results), then having this capacity amounts to having "perfect recall" (a standard assumption in epistemic logic). Further, we require the state Σ^t_O of the universe according to O (if such a state exists) to be at any given time t a separable state of the form $(\Sigma - O)_O \otimes O_O$, where $O_O = \bigotimes_{i \in O} s_i$ is a fully separable state. More generally, this poses certain constraints on the relative state of every subsystem $S \subseteq \Sigma$, namely the state S^t_O must at every time be of the form $(S - O)_O \otimes (S \cap O)_O$ (where $(S \cap O)_O$ is the restriction $\bigotimes_{i \in S \cap O} s_i$ to $S \cap O$ of the fully separable state O_O). These assumptions lead us to state the following framework in which the notions of history and protocol and admissible observer can be formally defined:

[1] The notion of admissible observer is applicable to any system, including a measurement apparatus as well as systems that are typically referred to as 'agents' in epistemic logic, AI and computer science.

Histories and protocols. A *history (of length n) of a system S wrt an observer O* is a sequence $h = (S_O^0, T^1, S_O^1, T^2, \ldots, S_O^{n-1}, T^n, S_O^n)$, where for all times $k < n$, S_O^k are state descriptions of the system S (with respect to O) s.t. $S_O^k = T^k(S_O^{k-1})$, and where all T^k are linear maps (called *dynamical maps*), requiring that T^k are either unitary maps on $H^{(S)}$ or they are 'projective' maps (of the form $P_{\rho_A} \otimes B_O^{\rho_A} \otimes I_{S-(O\cup A)}$, modelling an O-measurement of some subsystem $A \subseteq S - O$ that collapses the observed system A onto a pure state ρ_A, while recording it as $B_O^{\rho_A}$ into O's memory). Given any time moment $t \leq n$, the *t-restriction of the history $h = (S_O^0, T^1, \ldots, S_O^n)$* is the initial t-long subsequence $h|t := (S_O^0, T^1, \ldots, T^t, S_O^t)$.

A **protocol** (for S wrt an observer O) is a family π of histories of the same length n and having the same initial state S_O^0.

Admissible observers. A subsystem $A \subseteq S$ is an *admissible observer wrt (a background observer system) O and a given history/protocol* (for S wrt O) iff, for all that O knows (at any given time), *A may be informationally persistent* throughout the history/protocol; in other words, if O never gets to know for a fact that A is not (will not be, or has not been) persistent throughout the history/protocol.

Communities of observers. A group $\mathcal{G}$ of subsystems of a system S (that includes O itself) is a *community of admissible observers* wrt O (and a given history/protocol) if every system $A \in \mathcal{G}$ is an admissible observer both wrt to any other system $B \in \mathcal{G}$ as well as wrt O (and the given history/protocol).

With these definitions, we can state our main proposition:

Proposition. Given a protocol π for a system S wrt an observer O, let $A \subseteq S$ be an admissible observer. For every description of the state evolution of the system S wrt O (as defined above), there exists an alternative description of the evolution of S wrt A (in which A takes the place of the background observer), such that the two descriptions are 'compatible' in a weak sense, namely relative to the protocol π.

Here, the notion "compatibility" of state descriptions is weaker than compatibility in the sense of Leifer and Spekkens [23], and it is only *relative to the given protocol*: at each time t, the two state descriptions can be refined to states that *assign the same probabilities to all possible outcomes of all the measurements that can happen according to the protocol π*.

We refer to [6] for the proof-sketch of this result. The formalization uses input from epistemic logic, in particular to define the notion of propositional knowledge and to reason about nestings of knowledge expressions. Indeed, *only for a community of admissible observers we can meaningfully use all the power of epistemic logic*, with indefinitely long nested levels of knowledge, e.g. $K_O K_A K_B K_A K_C \varphi$ (where each knowledge operator is labeled by an observer). All the standard axioms of Epistemic Logic hold in this case, and so in particular the principle of 'Knowledge Transfer' (called condition (C) in [18]) holds (since it is a theorem of Epistemic Logic). As a consequence, the agents involved in any standard protocol studied in Quantum Information and Quantum Computation need always to be assumed to form a community of observers.

3 Conclusion

The view we adopt is that, on the one hand, every system can be treated as an 'observer' from its own perspective; but, on the other hand, *not* all external systems can be treated as "observers" from the perspective of a given observer. Instead of blankly stating that Quantum Mechanics does not allow observers to reason about other systems as observers, a more fine-grained solution is indeed possible. Being an observer is itself a relational property, one that is relative to another observer (or to a community of other observers) and a background protocol (that restricts the timeframe and the possible histories). This doesn't make observers "less real" and less interesting, and it doesn't make Epistemic Logic less useful for reasoning about quantum scenarios, on the contrary.[2]

[2]In this sense, see some of our previous work [1, 2, 3, 5, 4], on using epistemic logic to reason both about the informational properties of quantum systems (and in particular about epistemic characterizations of entanglement), and about the epistemic states of classical observers in quantum environments.

References

[1] Baltag, A.; Smets, S.: A dynamic-logical perspective on quantum behavior, Studia Logica, 89:185-209 (2008)

[2] Baltag, A.; Smets, S.: Logics of Informational Interactions, Journal of Philosophical Logic, 44(6):595-607 (2015)

[3] Baltag, A.; Smets, S.: Modeling correlated information change: from conditional beliefs to quantum conditionals, Soft Computing, 21(6), 1523-1535 (2017)

[4] Baltag, A.; Smets, S.: Correlated Knowledge, An Epistemic-Logic View on Quantum Entanglement, International Journal of Theoretical Physics, 49:3005-3021 (2010)

[5] Baltag, A.; Smets, S.: Correlated Information: A Logic for Multi- Partite Quantum Systems. In Coecke, B. and Panangaden, P. (eds.) Electronic Notes in Theoretical Computer Science ENTCS. Proceedings of the 6th Workshop on Quantum Physics and Logic, Oxford. vol. 270, pp.3-14 (2011)

[6] Baltag, A.; Smets, S.: Logic Meets Wigners Friend (and their Friends). Int J Theor Phys 63, 97 (2024).

[7] Barwise, J.; Perry, J.: Situations and Attitudes, Bradford Books, MIT Press (1983)

[8] Boge, F.: Quantum information vs. epistemic logic: An analysis of the Frauchiger-Renner theorem. Foundations of Physics, 49:1143-1165, arXiv:1909.11889v1 (2019)

[9] Brukner, C.: On the quantum measurement problem, in R. Bertlmann and A. Zeilinger (eds): Quantum [Un]speakables II, Half a Century of Bell's Theorem. Springer (2017), arXiv:1507.05255v1 (2015)

[10] Brukner, C.: A no-go theorem for observer-independent facts, Entropy 20, 350 (2018)

[11] Guérin, P.A.; Baumann, V.; Del Santo, F.; Brukner, C.: A no-go theorem for the persistent reality of Wigner's friend's perception. Commun Phys 4, 93 (2021).

[12] DeBrota, J.B; Fuchs, C.A; Schack, R.: Respecting One's Fellow: QBism's Analysis of Wigner's Friend, Foundations of Physics 50:1859-1874 (2020)

[13] Di Biagio, A.; Rovelli, C.: Stable Facts, Relative Facts, Foundations of Physics 51:30 (2021)

[14] Corti, A.; Fano, V.; Tarozzi G.,: A Logico-Epistemic Investigation of Frauchiger and Renner's Paradox, International Journal of Theoretical Physics 62:54, (2023)

[15] Deutsch, D.: Quantum theory as a universal physical theory, International Journal of Theoretical Physics, 24, I (1985)

[16] Dretske, F.: Knowledge and the Flow of Information, MIT Press (1981)

[17] Elouard, C.; Lewalle, P.; Manikandan, S.K; Rogers, S; Frank, A. and Jordan, A.N: Quantum erasing the memory of Wigner's friend. Quantum, 2021-07-08, volume 5, page 498, arXiv:2009.09905v4 (2021)

[18] Frauchiger, D.; Renner, R.: Quantum theory cannot consistently describe the use of itself, Nature Communications, 9(1):3711. arXiv:1604.07422 (2018)

[19] Fagin, R.; Halpern, J.; Moses, Y. and Vardi, M.: Reasoning about Knowledge, MIT Press (1995)

[20] Healey, R.: Quantum Theory and the Limits of Objectivity Richard, Foundations of Physics, 48:1568-1589, (2018)

[21] Laudisa, F. and Rovelli, C.: Relational Quantum Mechanics, E. N. Zalta (ed.), The Stanford Encyclopedia of Philosophy, online (2021)

[22] Laloë, F.: Can quantum mechanics be considered consistent? a discussion of Frauchiger and Renner's argument. ArXiv:1802.06396v3 (2018)

[23] Leifer, M.S.; Spekkens, R.W.: A Bayesian approach to compativility, improvement and pooling of quantum states, Journal of Physics A: Mathematical and Theorical, 47, 275301, arXiv:1110.1085v1 (2014)

[24] Nielsen, M. A.; Chuang, I.L.: Quantum Information and Quantum Computation. Massachusetts Institute of Technology (2010)

[25] Nurgalieva, N.; Renner, R.: Testing Quantum Theory with Thought Experiments, Contemporary Physics, 61(3):193-216, arXiv:2106.05314v1 (2020)

[26] Nurgalieva, N.; del Rio, L.: Inadequacy of Modal Logic in Quantum Settings, EPTCS 287:267-297 (2019)

[27] Rovelli, C.: Relational Quantum Mechanics, International Journal of Theoretical Physics, 35(8):1637-1678, (1996)

[28] Adlam, E.; Rovelli, C. :Information is Physical: Cross-Perspective Links in Relational Quantum Mechanics, in Philosophy of Physics 1(1): 4, 1-19 (2023)

[29] Shimony, A.: Role of the Observer in Quantum Theory, American Journal of Physics, 31:755 (1963)

[30] van Ditmarsch, H.P; van der Hoek, W.; Kooi, B.: Dynamic Epistemic Logic, Springer (2007)

[31] Wigner, E.P.: Remarks on the mind-body question, in I.J. Good, The Scientist Speculates, London Heinemann (1961)

[32] Waaijer, M.; van Neerven, J.: Relational Analysis of the Frauchiger-Renner Paradox and Interaction-Free Detection of Records from the Past, Foundations of Physics (2021)

[33] Crull, E. and Bacciagaluppi, G.: Translation of W. Heisenberg: "Ist eine deterministische Ergänzung der Quantenmechanik möglich?", Pittsburgh Phil-Sci Archive, (2011)

[34] von Neumann, J.: Mathematical Foundations of Quantum Mechanics , N. Wheeler (ed.) of the English translation, New Edition, 2018, Princeton University Press. First published in German in (1932)