Impossible worlds, triviality, and duality

Jens Christian Bjerring
Aarhus University
jcbjerring@gmail.com

May 2014

Introduction

- In a letter to Graham Priest, David Lewis writes:

  > *I'm increasingly convinced that I can and do reason about impossible situations. [...] But I don't really understand how that works. Paraconsistent logic as developed by you [Priest] and your allies is clear enough, but I find it a bit off the topic. For it allows (a limited amount of) reasoning about blatantly impossible situations. Whereas what I find myself doing is reasoning about subtly impossible situations, and rejecting suppositions that lead fairly to blatant impossibilities. (Lewis 2004, pp. 176-177.)*

- "Lewisian challenge": to model the doxastic lives of minimally rational agents like Lewis, find a model of belief that ensures that only possible and subtly impossible worlds remain doxastically possible for these agents.
  - Because minimally rational, they can immediately rule out as doxastically possible *blatant* logical inconsistencies such as $\{A, \neg A\}$ and $\{\neg A, (A \wedge B)\}$, but not necessarily *subtle* logical inconsistencies such as $\{\neg(\neg(A \to C) \to \neg((\neg A \to \neg B) \to \neg(\neg A \to \neg C)))\}$.
- Claim: we cannot meet the Lewisian challenge in an impossible-worlds framework for doxastic possibility and necessity.

Beliefs and doxastic necessity

- What is a world-involving framework for belief?

  (**Belief**) An agent *a* believes *A* (*A* is doxastically possible for *a*) iff *A* is true at all worlds that are doxastically possible for *a*.

- If all doxastically possible worlds are logically possible, agents are modelled as logically omniscient: they believe all logical consequence of what they believe (including all logical truths).
  - Suppose *a* believes *A*. Then *A* is true at all (logically possible) worlds that are doxastically possible for *a*. Consider any *B* that follows logically from *A*. Then *B* is true at all worlds that are doxastically possible for *a*. So *a* believes *B*.
- To model logically non-omniscient agents, we appeal to *logically impossible worlds* where the truths of logic can fail.
  - Suppose *a* believes *A*. Then *A* is true at all worlds that are doxastically possible for *a*. Consider any *B* that follows logically from *A*. Since some impossible worlds might remain doxastically possible for *a*, *B* need not be true at all worlds that are doxastically possible for *a*. So *a* need not believe *B*.

Impossible worlds

- What are impossible worlds?
  - Possible worlds = maximal, consistent sets of sentences in $\mathcal{L}$.
  - Impossible worlds = maximal, inconsistent sets of sentences in $\mathcal{L}$.
    - Maximality: a set Γ of sentences is *maximal* iff for all sentences A, either $A \in \Gamma$ or $\neg A \in \Gamma$.
    - Consistency: a set Γ of sentences is *consistent* iff Γ is satisfiable, where Γ is *satisfiable* iff there is a propositional evaluation that makes all sentences in Γ true.
    - $\mathcal{L}$: some sufficiently strong and precise world-making language that contains all sentence types in English and symbols $\neg$ and $\wedge$ that play the same inferential roles as classical negation and conjunction.
- What does it mean for a sentence to be true at a world?

  (**Truth**) A is *true* at a world w iff $A \in w$.
  (**Falsity**) A is *false* at a world w iff $A \notin w$.

    - In turn, we can also say that a world is maximal whenever for any A, either A is true at w or $\neg A$ is true at w.
- To address the Lewisian challenge, we now need to know what distinguishes blatantly from subtly impossible worlds.

Blatantly impossible worlds

- Intuitively, an inconsistency is blatantly inconsistent whenever it can be ruled out easily, effortlessly, or immediately by a minimally rational agent.
  - A minimally rational agent is *minimally logical competent* in the following sense: she never fails to believe any sentence that she can infer from what she already believes by 1 step of logical reasoning.
- We can use this idea to characterize blatant inconsistencies:

  (**Blatant**) A set Γ of sentences is *blatantly inconsistent* just in case a contradiction $\{A, \neg A\}$ can be inferred from Γ by use of at most 1 application of an inference rule in a standard proof system for propositional logic; otherwise, if inconsistent, Γ is *subtly inconsistent*.

  - $\{A, \neg A\}$ and $\{\neg(A \wedge B), A, B\}$ are blatantly inconsistent, whereas $\{\neg(\neg(A \rightarrow C) \rightarrow \neg((\neg A \rightarrow \neg B) \rightarrow \neg(\neg A \rightarrow \neg C)))\}$ is only subtly inconsistent.
- Blatantly impossible worlds are thus worlds that contain a blatant inconsistency.
- Despite the weak characterization of the current impossible-worlds framework, we can show that it cannot be used to meet the Lewisian challenge.

Impossibility result

- For the following result can be shown:

  (**Incon**) All maximal, logically inconsistent sets of sentences contain an instance of a LNC-, CF-, or NCF-inconsistency:

  LNC-inconsistency (law of non-contradiction): $\{A, \neg A\}$.
  CF-inconsistency (conjunction fallacy): $\{\neg A, (A \wedge B)\}$, $\{\neg B, (A \wedge B)\}$, $\{\neg A, \neg B, (A \wedge B)\}$.
  NCF-inconsistency (negated conjunction fallacy): $\{\neg(A \wedge B), A, B\}$.

- Given (Blatant), we have:

  (**Incon**) All logically impossible worlds are blatantly impossible.

  - That is: all maximal, inconsistent sets contain a set from which $\{A, \neg A\}$ can be inferred in 1 step of logical reasoning by use of simple propositional rules.

- In turn, the following result obtains:

  (**Result**) There is no modal space such that:

  (R_1) there are impossible worlds;
  (R_2) there are no non-maximal worlds;
  (R_3) there are no blatantly impossible worlds.

- So, given (Result), it seems that we cannot use the impossible-worlds framework to meet the Lewisian challenge.

Problem and task

- But (Result) loses its *formal* bite with respect to (Belief) when the doxastic $\Box$ and $\Diamond$ are duals: $\Diamond A \leftrightarrow \neg\Box\neg A$.[1]

  $\Box A$: A is doxastically necessary for a (a believes A).
  $\Diamond A$: A is doxastically possible for a.

- Reason: (Pos) fails when impossible worlds are allowed.

  (**Nec**) $\Box A$ iff A is true at all worlds that are doxastically possible for a.
  (**Pos**) $\Diamond A$ iff A is true at some world that is doxastically possible for a.

  - Counterexample: Let $\{w\}$ be the set of worlds that are doxastically possible for a, and suppose that both A and $\neg A$ are true at w. Then, by (Nec), $\Box A$, and, by duality, $\neg\Diamond\neg A$. However, by (Pos), right to left, $\Diamond\neg A$. So contradiction. So (Pos) fails.

- If (Pos) fails, (Result) need not worry us: even if some blatantly impossible worlds are doxastically possible for a minimally rational agent, it need not mean that any blatant inconsistencies are.
- Task: to establish the formal force of (Result), we need a version of (Pos).

[1] As pointed out to me by Kit Fine.

Acceptance and Rejection 1

- In the context of non-ideal agents who may believe both A and $\neg A$, it is reasonable to give up the duality of $\Box$ and $\Diamond$.
  - Constraint: If a believes A, A should count as doxastically possible for a.
  - Suppose $\Box A$ and $\Box\neg A$. When $\Diamond$ is the dual of $\Box$, it follows that $\neg\Diamond\neg A$. Yet, when $\Box\neg A$, it should follow that $\Diamond\neg A$. So $\Diamond$ and $\Box$ cannot be duals.
- Instead: interpret doxastic necessity and possibility in terms of the primitive notions of *acceptance* and *rejection* of sets of sentences.
  - Similar pairs of attitudes in the literature: belief - disbelief (Russell 1913); belief - refusal to believe (Priest 2005 and 2006); assertion - denial (Restall 2005); hooraying - booing (Blackburn 1984); “Yes” and “No” (Rumfitt 2000).
  - Intuitively, to accept Γ is to answer “yes” to the question “Does Γ describe a way the world is?”
  - Intuitively, to reject Γ is to answer “no” to the question “Does Γ describe a way the world is?”
- Central constraints on acceptance and rejection:
  - No agent can simultaneously accept and reject Γ.
  - Acceptance and negation is conceptually distinct from rejection.
    - For instance: an agent can accept $\{\neg A\}$ without rejecting $\{A\}$.

Acceptance and rejection 2

- We can then define doxastic necessity, impossibility, and possibility:

(**DN**) Γ is doxastically necessary for an agent *a* iff *a* accepts Γ.
(**DI**) Γ is doxastically impossible for an agent *a* iff *a* rejects Γ.
(**DP**) Γ is doxastically possible for an agent *a* iff *a* does not reject Γ.

- Next we can introduce three predicates NEC, IMP, POS in the metalanguage:

NEC(Γ) = Γ is doxastically necessary for *a*.
IMP(Γ) = Γ is doxastically impossible for *a*.
POS(Γ) = not: IMP(Γ) = Γ is doxastically possible for *a*.

- Our task is thus to establish the following versions of (Nec) and (Pos):

(**Dox-Nec**) NEC(Γ) iff for all *w* that are doxastically possible for *a*, each $A \in \Gamma$ is true at *w*.
(**Dox-Pos**) POS(Γ) iff for some *w* that is doxastically possible for *a*, each $A \in \Gamma$ is true at *w*.

Doxastic possibility: worlds

- To prove (Dox-Nec) and (Dox-Pos), let us define:

  (**Dox-Pos-**w) A set W_a of worlds is doxastically possible for an agent a iff:

  (i) If NEC(Γ), then it is case that for all $w \in W_a$: $\Gamma \subseteq w$.
  (ii) If IMP(Γ), then it is case that for all $w \in W_a$: $\Gamma \not\subseteq w$.

- So a world is doxastically possible for an agent when it contains everything she accepts and nothing she rejects.
- Using Lindenbaum-style reasoning, we can establish (iii) and (iv) from (Dox-Pos-w):

  (iii) If not NEC(Γ), then it is the case for some $w \in W_a$: $\Gamma \not\subseteq w$.

  (iv) If not IMP(Γ), then it is the case for some $w \in W_a$: $\Gamma \subseteq w$.

  - Roughly, for (iii): If not NEC(Γ), then there is at least one $\{B\} \subseteq \Gamma$ such that not NEC($\{B\}$). Let Ax be the set of accepted sentences. For each A—except B—define a rule that adds A to Ax just in case IMP does not apply to the resulting set $Ax \cup \{A\}$; otherwise it adds $\neg A$ to Ax.
  - Eventually, either A or $\neg A$—or both—will be added to Ax, and the resulting set will correspond to a maximal world w such that $\Gamma \not\subseteq w$ and such that w satisfies (i) and (ii) in (Dox-Pos-w).

Proof of (Dox-Nec)

(**Dox-Nec**) NEC(Γ) iff for all w that are doxastically possible for a, each $A \in \Gamma$ is true at w.

- Left to right: Assume NEC(Γ). By (Dox-Pos-w), (i), then $\Gamma \subseteq w$ for all w that are doxastically possible for a. In turn, for all worlds w doxastically possible for a, $A \in w$ for each $A \in \Gamma$. So, by (Truth), for all worlds w that are doxastically possible for a, each $A \in \Gamma$ is true at w.
- Right to left: Assume, for all worlds w that are doxastically possible for a, that each $A \in \Gamma$ is true at w. Given (Truth), then $A \in w$ for each $A \in \Gamma$, in which case $\Gamma \subseteq w$ for all w that are doxastically possible for a. Suppose, for reductio, that not NEC(Γ). Then, by the derived condition (iii), $\Gamma \not\subseteq w$ for some w that is doxastically possible for a. Contradiction. So NEC(Γ).

- So (Doc-Nec) holds.

Proof of (Dox-Pos)

(**Dox-Pos**) POS(Γ) iff for some w that is doxastically possible for a, each $A \in \Gamma$ is true at w

- Left to right: Assume POS(Γ). Then not IMP(Γ). Then, by the derived (iv), there is some world w that is doxastically possible for a such that $\Gamma \subseteq w$. Then $A \in w$ for each $A \in \Gamma$, in which case, by (Truth), each $A \in \Gamma$ is true at w for some w that is doxastically possible for a.
- Right to left: Assume, for some world w that is doxastically possible for a, that each $A \in \Gamma$ is true at w. Given (Truth), then $A \in w$ for each $A \in \Gamma$, in which case $\Gamma \subseteq w$ for some w that is doxastically possible for a. Suppose, for reductio, that IMP(Γ). Then, by (Dox-Pos-w), condition (ii), $\Gamma \nsubseteq w$ for all w that are doxastically possible for a. Contradiction. So not IMP(Γ). So POS(Γ).

- So (Dox-Pos) holds.

Conclusion

- Given (Dox-Pos), we have restored the formal force of (Result): since some blatantly impossible worlds must remain doxastically possible for a minimally rational agent, some blatant inconsistencies must remain doxastically possible for these agents.
- But this is not plausible.
- So when worlds are maximal, I conclude that we cannot use the impossible-worlds framework to meet the Lewisian challenge.
    - In fact, although I cannot argue for it here, whether or not worlds are maximal, I claim that we cannot use an impossible-worlds framework to model the doxastic lives of minimally rational agents like David Lewis.

Literature

Bjerring, J. C. (2013). Impossible worlds and logical omniscience: an impossibility result. In *Synthese* 190, pp. 2505-2524.

Blackburn, S. (1984). *Spreading the Word*. Oxford: Oxford University Press.

Lewis, D. (2004). Letters to Priest and Beall. In B. Armour-Garb, J.C. Beall, G. Priest (Eds.), *The Law of Non-Contradiction—New Philosophical Essays* (pp. 176-177). Oxford: Oxford University Press.

Priest, G. (2005). *Towards Non-Being: the Logic and Metaphysics of Intentionality*. Oxford: Oxford University Press.

Priest, G. (2006). *Doubt Truth to be a Liar*. Oxford: Oxford University Press.

Restall, G. (2005). Multiple Conclusions. In *Logic, Methodology and Philosophy of Science: Proceedings of the Twelfth International Congress* (Hajek, P., Valdes-Villanueva, L., Westerstahl, D. (eds.)). Kings College Publications (2005), pp. 189-205.

Rumfitt, I. (2000). “Yes” and “No”, in *Mind*, 109, pp. 781-823.

Russell, B. (1913/1992). *Theory of Knowledge: The 1913 Manuscript*. New York: Routledge.