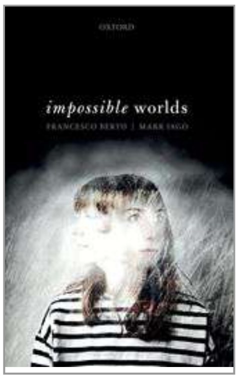


Impossible Worlds
Francesco Berto and Mark Jago

Print publication date: 2019
Print ISBN-13: 9780198812791
Published to Oxford Scholarship Online: August 2019
DOI: 10.1093/oso/9780198812791.001.0001

The Logic of Imagination

Francesco Berto
Mark Jago

DOI:10.1093/oso/9780198812791.003.0007

Abstract and Keywords

Imagination seems to have a logic, albeit one which is hyperintensional and sensitive to context. This chapter offers a semantics of imagination, with operators expressing 'imaginative acts' of mental simulation. A number of conditions that could be imposed on the semantics are then discussed, in order to validate certain inferences. One important issue is how acts of imagination interact with disjunction: one can imagine some disjunction as obtaining without being imaginatively specific about which disjunction obtains. This chapter subsequently turns to non-monotonicity: how B may follow from imagining that A, but not from imagining that $A \wedge C$. Finally, the *Principle of Imaginative Equivalents* is discussed.

Keywords: imagination, semantics of imagination, mental simulation, non-monotonicity, Principle of Imaginative Equivalents

7.1 Hyperintensional Imagination

'Imagining' is highly ambiguous, as we saw in §1.5. We use the word for such different mental activities as daydreaming, hallucinating, supposing, planning, make-believing. We will focus on mental states with a propositional content (imagining *that A*: that Obama is blond-haired, that Holmes walks through Victorian London). We will target a notion found in such widely discussed works on imagination as Chalmers (2002b) and Yablo (1993), and dubbed by the latter 'positive conceivability'.

Positively conceiving that A is understood as a mental operation different from merely supposing or assuming that A, as when we make an assumption in a mathematical proof and, in some sense, as more substantive (Balcerak Jackson 2016). We represent a situation – a configuration of objects and properties – of which A is a truthful description:

> Positive notions of conceivability require that one can form some sort of positive conception of a situation in which A is the case. One can place the varieties of positive conceivability under the broad rubric of *imagination*: to positively conceive of a situation is to imagine (in some sense) a specific configuration of objects and properties. … Overall, we can say that A is positively conceivable when one can imagine that A: that is, when one can imagine a situation that verifies A.
>
> (Chalmers 2002b, 150, notation modified)

(p.142) Similarly, Yablo (1993) speaks of conceiving that A as imagining a *world* verifying A. (Yablo grants that we do not imagine the world in all detail.) This seems to be the notion typically at issue in debates on whether conceivability entails possibility (see e.g. Gendler 2000, Hill 1997, Kung 2010, Roca-Royes 2011, Stoljar 2007, and the essays in Gendler and Hawthorne 2002). As we also saw in §1.5, in these debates it is often not clear what kind of mental representation is involved in the relevant act of conceiving or imagining. It is not clear whether it involves linguistic mental representations, or pictorial mental imagery mimicking corresponding sensory modalities. (We'll return to the issue in §7.3.)

Conceivability in the sense of Chalmers and Yablo seems to be linked to mental simulation, a phenomenon studied in cognitive psychology. We simulate alternatives to reality in our mind, in order to explore what would and would not happen if they were realized. This helps us to cope with reality itself, by improving future performance and allowing us to make contingency plans. (The works in Markman et al. 2009 explore various empirical phenomena in this ballpark.) That some things would happen in the envisaged scenario, and some would not, seems to imply that such exercises have some kind of logic: some things follow in the imagined situation, and some do not (Byrne 2005).

Works on the logic of imagination typically resort to a possible worlds framework, modelling imagination as a restricted quantifier over possible worlds (Costa Leite 2010, Niiniluoto 1985). But imagination, *qua* intentional mental state, is hyperintensional. Lois Lane can imagine that Superman is in love with her without imagining that Clark Kent is in love with her, as she ignores their being identical; we can imagine proving that 107 + 215 = 322 without imagining proving Fermat's Last Theorem; and we can imagine that water turns out not to be H_2O (§1.5). This makes the phenomenon difficult to model via standard possible worlds semantics.

Wansing (2017) uses *neighbourhood semantics*, which we met in §5.2, for his logic of imagination. This allows several logical closure properties to fail for it: one's imagining that *A*, and that if *A* then *B*, **(p.143)** for instance, does not entail one's imagining that *B*. However, it is still the case that, as a consequence of the adoption of neighbourhood semantics, if *A* is logically equivalent to *B* and one imagines that *A*, one automatically imagines that *B* and vice versa. But this result seems wrong. Even in weak logics such as the basic relevant logic **B** of §6.1, *A* is equivalent to $A \vee (A \wedge B)$. However, one may imagine that *A* without imagining that $A \vee (A \wedge B)$, for every *B*. One may lack the very concepts involved in *B*, for example.

Impossible worlds are thus natural candidates for modelling imagination as mental simulation. But imagination, so understood, seems to have further features with which an acceptable model must comply. We'll describe them in the rest of this section, and present a simple impossible worlds semantics in the following §7.2, drawing on Berto (2014, 2017).

One feature of imagination as mental simulation is that it can be voluntary in ways belief cannot. One can imagine that all of one's home town has been painted yellow but, having overwhelming evidence of the contrary, one cannot easily make oneself believe it. Conscious acts of imagination as mental simulation can have an arbitrary, explicit starting point (Langland-Hassan 2016, Wansing 2017). This may be determined by the agent (as in, 'now let's imagine what would happen if …'), or it may be helped by external inputs (think of going through a novel, taking the sentences you read as your explicit input). According to Nichols and Stich's influential *mental simulation model* (2003), we begin imagining with 'an initial premiss or set of premisses, which are the basic assumptions about what is to be pretended' (2003, 24).

Imagination is not purely inferential, however. 'Children and adults elaborate the pretend scenarios in ways that are not inferential', filling in the explicit instruction with 'an increasingly detailed description of what the world would be like if the initiating representation were true' (Nichols and Stich 2003, 26–8; see also Langland-Hassan 2016, van Leeuwen 2016). You read a Jeffery Deaver book featuring Lincoln Rhyme, a detective working in New York on some murder case. The sentences of the book give you the explicit input. You integrate it **(p.144)** with background information you've imported into the scenario, on the basis of what you know or believe: New York is in the US, and normally detectives are human beings, although (let's suppose) the Deaver story does not state these things explicitly. Absent information to the contrary, you imagine Lincoln as a human being working in the US, although this is not entailed logically by the explicit input.

We propose to model this via modal operators interpreted as *variably* strict quantifiers over worlds, possible or otherwise. The variability of the quantifiers accounts for the contextual selection of the information we import in acts of imagination. As we will see, the input will play a role similar to a conditional's antecedent in Lewis's (1973b) semantics for counterfactuals.

It's important, however, not to treat agents as importing too much background information into acts of imagination. We do not indiscriminately import arbitrary, unrelated contents into imagined scenarios. You know that Manila is the capital of the Philippines, but this is immaterial to your imagining Lincoln Rhyme's New York adventures. Such adventures do not involve Manila or the Philippines at all. So you will not, in general, import such irrelevant content in your scenario. Of course, you *can* imagine things about Manila as well, by some free-floating association of ideas; but you will avoid it while engaging in mental simulation specifically of Lincoln Rhyme's New York adventures. So such exercises of imagination must obey some constraint of relevance.

7.2 A Semantics of Imagination

We will use a propositional language $\mathscr{L}$ with the usual set of atoms AT closed under negation ¬, conjunction ∧, disjunction ∨, a strict conditional ⥽, modal operators □ and ◇, and square brackets '[' and ']', put to special use. The well-formed formulas are the atoms and, if A and B are well-formed formulas, then so are:

$$\neg A \mid (A \wedge B) \mid (A \vee B) \mid (A \rightarrowtail B) \mid \Box B \mid \Diamond B \mid [A]B$$

(p.145) Things of the form '$[A]$' are modal operators indexed by formulas. (In conditional logic, this idea goes back to Chellas (1975).) Take a bunch of acts of imagination, performed by a given agent on specific occasions, and characterized by an explicit input: what the agent sets out to imagine ('Let's imagine that Holmes chases Moriarty across London in a horse-drawn carriage'), which can be taken as corresponding to Nichols and Stich's 'initial premiss', the 'basic assumption about what is to be pretended'. This is given directly by a formula of $\mathscr{L}$. If K is the set of formulas expressing possible explicit inputs, then for each $A \in K$, $[A]$ is the corresponding modal. (K might be the whole language, or the language free of the '$[A]$' operators, or some restricted fragment of it. Just which restrictions should be put on K is an interesting issue, which we will not pursue here.) We can read '$[A]B$' as 'it is imagined in the act with explicit input A, that B'; or, more tersely, 'it is imagined in act A that B'. We will call each $[A]$ an *imagination operator*.

The semantics is inspired by the relational frames for **FDE** (§5.4). An *imaginative frame* $\mathscr{F}$ is a triple $\langle\langle W, N, \{R_A \mid A \in K\}\rangle\rangle$. W is the set of worlds; $N \subseteq W$ is the subset of normal worlds; the worlds in $W - N$ are the non-normal or impossible worlds; and each $R_A \subseteq W \times W$ is a binary accessibility relation on W, one for each sentence $A \in K$.

A frame becomes a model $\mathscr{M} = \langle\langle W, N, \{R_A | A \in K\}, \rho \rangle\rangle$ when endowed with a valuation relation ρ, relating (for each world w) the atoms in AT to truth ('$\rho_w p1$'), falsity, ('$\rho_w p0$'), both, or neither. We then extend ρ to the whole language as follows. For the extensional connectives we have, for all $w \in N$:

(S1¬) $\rho_w(\neg A)1$ iff $\rho_w A0$

(S2¬) $\rho_w(\neg A)0$ iff $\rho_w A1$

(S1∧) $\rho_w(A \wedge B)1$ iff $\rho_w A1$ and $\rho_w B1$

(S2∧) $\rho_w(A \wedge B)0$ iff $\rho_w A0$ or $\rho_w B0$

(S1∨) $\rho_w(A \vee B)1$ iff $\rho_w A1$ or $\rho_w B1$

(S2∨) $\rho_w(A \vee B)0$ iff $\rho_w B0$ and $\rho_w B0$

(p.146) The familiar modalities get their usual **S5** clauses, over normal worlds. For all $w \in N$:

(S1⥽) $\rho_w(A ⥽ B)1$ iff for all $w_1 \in N$, if $\rho_{w_1} A1$, then $\rho_{w_1} B1$

(S2⥽) $\rho_w(A ⥽ B)0$ iff for some $w_1 \in N$, $\rho_{w_1} A1$, and $\rho_{w_1} B0$

(S1□) $\rho_w(\Box A)1$ iff for all $w_1 \in N$, $\rho_{w_1} A1$

(S2□) $\rho_w(\Box A)0$ iff for some $w_1 \in N$, $\rho_w A0$

(S1◇) $\rho_w(\Diamond A)1$ iff for some $w_1 \in N$, $\rho_{w_1} A1$

(S2◇) $\rho_w(\Diamond A)0$ iff for all $w_1 \in N$, $\rho_w A0$

As for the $[A]$s, for $w \in N$:

(S1[A]) $\rho_w([A]B)1$ iff for all $w_1 \in W$ such that $R_A ww_1$, $\rho_{w_1} B1$

(S2[A]) $\rho_w([A]B)0$ iff for some $w_1 \in W$ such that $R_A ww_1$, $\rho_{w_1} B0$

Read '$R_A ww_1$' as saying that w_1 is accessed by an act of imagination with explicit input A, performed at w.

These recursive truth conditions have been defined for worlds in N. For worlds in $W - N$, ρ relates all formulas directly to truth values, irrespective of their syntax. (We met this approach to non-normal worlds in Rantala frames for epistemic logic, §5.3.) Logical validity and consequence are defined, once again, as truth and truth preservation (respectively) at all normal worlds in all models.

At normal worlds, the recursive truth and falsity conditions for $[A]$-formulas can also be expressed using set-selection functions, as in Lewis's (1973b) semantics for counterfactuals. Each $A \in K$ comes with a projection function f_A, taking as input the world where the act is performed and giving the set of worlds made accessible: $f_A w = \{w_1 \in W \mid R_A w w_1\}$.

A key idea in Lewis's semantics is that f_A outputs the A-worlds that are most objectively similar to the input world. In the context of a semantics for imagination, we can take it as outputting the worlds **(p.147)** that are most subjectively plausible for the agent, given input A. Imagination, on this model, works as a kind of belief revision. We will not impose a relation of comparative plausibility for worlds, as Baltag and Smets (2006), Grove (1988), and Leitgeb and Segerberg (2007) do. How orderings of this kind should work when impossible worlds are around involves a number of open issues, and the semantics below is largely independent of them. We will come back to plausibility orderings involving impossible worlds in §11.3, in the context of a treatment of truth in fiction.

Let $|A|$ be the set of worlds where A is true. Then, for $w \in N$:

(S1$[A]$) $\rho_w([A]B)1$ iff $f_A w \subseteq |B|$

(S2$[A]$) $\rho_w([A]B)0$ iff $f_A w \cap |\neg B| \neq \emptyset$

So for normal worlds, $[A]B$ is true (false) at w iff B is true at all worlds (false at some world) in a set selected by f_A. The relational and functional clauses are equivalent, given that $R_A w w_1$ iff $w_1 \in f_A w$. But it will sometimes be easier to talk using one formulation rather than the other.

A natural constraint on the semantics is that, for all $A \in K$ and $w \in N$:

(OBTAINING) If $w \in N$, then $f_A w \subseteq |A|$

Normal worlds access only those worlds where the explicit imaginative input *obtains*. Thus, in an act of imagination with explicit input A, one looks only at worlds where A is true. (We restrict this to normal worlds, since the non-normal/impossible worlds can do what they like.) We will consider only models satisfying OBTAINING.

To represent the imaginability of inconsistencies, we have allowed formulas to be related both to truth and to falsity (or to neither). But we may not want this to happen at normal worlds. We can add to the semantics a Classicality Condition as we did for relevant logics in §6.1, requiring all normal worlds to be maximally consistent with respect to atoms:

(p.148)

(CC′) If $w \in N$, then for each $p \in AT$, either $\rho_w p1$ or $\rho_w p0$ but not both.

This generalizes to all formulas not including imagination operators (as an easy induction on the complexity of formulas shows). To extend it to the whole language, we would need a different falsity condition (S2[A]). One option is to take $[A]B$ to be false when it is not true: for $w \in N$,

(S2[A]′) $\rho_w([A]B)0$ iff it is not the case that $\rho_w([A]B)1$.

This prevents inconsistencies accessible via imagination in non-normal worlds from creating inconsistencies at normal worlds.

Once the framework has been ‘classicalized’ in this way, the logic induced by the semantics for the connectives other than the imagination operators is just normal propositional **S5**. Let us now move on to an exploration of what is and what isn’t valid in the semantics.

7.3 The Mereology of Imagination

Our first logical validity is:

$$\models [A]A$$

The explicit input is always imagined. This is immediately guaranteed by OBTAINING. Next come validities that involve conjunction. Firstly, for reasons to be discussed soon, we may want that, when one imagines that a conjunction is the case, one imagines each conjunct. The following condition does the trick:

(SIMPLIFICATION) For all $w \in N$: if $R_A w w_1$, and $\rho_{w_1}(B \wedge C)1$, then $\rho_{w_1} B1$ and $\rho_{w_1} C1$

This gives the following validities:

$$[A](B \wedge C) \models [A]B \qquad [A](B \wedge C) \models [A]C$$

(p.149) To see why these follow, suppose $w \in N$ and $\rho_w([A](B \wedge C))1$. By (S1[A]), for all R_A-accessible worlds w_1, $\rho_{w_1} B \wedge C1$. Then SIMPLIFICATION gives us $\rho_{w_1} B1$ and $\rho_{w_1} C1$ and so, by (S1[A]) again, we have $\rho_w[A]B1$ and $\rho_w[A]C1$.

The companion constraint to SIMPLIFICATION is:

(ADJUNCTION) For all $w \in N$: if $R_A ww_1$, $\rho_{w_1} B1$, and $\rho_{w_1} C1$, then $\rho_{w_1}(B \wedge C)1$

This gives us:

$$[A]B, [A]C \models [A](B \wedge C)$$

To see why, suppose $w \in N$, $\rho_w([A]B)1$ and $\rho_w([A]C)1$. By (S1[A]), for all R_A-accessible worlds w_1, we have $\rho_{w_1} B1$ and $C\rho_{w_1} C1$. By ADJUNCTION, $\rho_{w_1}(B \wedge C)1$ for any such w_1 and so, by (S1[A]), $\rho_w([A](B \wedge C))1$.

Imaginative accessibility is then *de facto* limited by SIMPLIFICATION and ADJUNCTION to worlds that behave with respect to conjunction. But ADJUNCTION may be found problematic. Is it so that, when one imagines that B and that C in a single act [A], one automatically imagines their conjunction?

A similar question has been asked for counterfactuals, namely whether different counterfactuals with the same antecedent demand the conjunction of their consequents, given the role consequents play in fixing the context of evaluation. Suppose one sets out to imagine Caesar, the Roman emperor, being in command of the US troops in the Korean War. (The example is Quine's (1960, 222).) This gives an explicit input, A. One can then unfold the scenario as one where Caesar uses the atom bomb, [A]B, if one imports into the representation information concerning the weapons available in the twentieth century. Or, one can get to imagine him using catapults, [A]C, if one rather allows the Roman military apparatus to step in. However, we shouldn't thereby infer [A]($B \wedge C$), involving Caesar's employing both bombs and catapults. One *can* imagine that, too, if one likes, but it should not come as an automatic logical entailment.

(p.150) We think that something has gone wrong with the reconstruction of the situation. Acts of imagination are contextually determined, in that the same explicit input can trigger the importation of different background information in different contexts (the time and place where the act takes place, the status of the agent's background information, and so on). In Quine's example, there is a clear contextual shift. So ADJUNCTION can be maintained by adding a contextual parameter. This can be represented in the formalism by a set of contexts and variables ranging over them, indexing imaginative acts. $[A]_x$ and $[A]_y$ will then stand for two distinct acts with the same explicit input A. Once the adjunctive inference is parameterized to contents with the same index, it avoids the worry in Quine's example.

We think ADJUNCTION is important in modelling an independently plausible conception of imagination. Imagination, in the relevant sense, is more than mere supposition of a content, A. Rather, it is positive conceivability in Chalmers's (2002b) sense: someone's representing a scenario or a state of affairs which makes A true. It is a generally agreed principle of the logic of (inexact) truthmaking (Fine 2014, Yablo 2014) that truthmakers behave in such conjunctive fashion. This mirrors the idea that states of affairs themselves can stand in parthood relations. (A conjunctive state of affairs that object m is P and object o is Q includes as constituents the individual states of affairs that m is P and that o is Q. We discuss logically complex states of affairs in Barker and Jago 2012 and Jago 2011.)

If imagination crucially involves mental imagery (as Kind (2001) argues), then there may be a certain mereological structure implemented in the mind, and we may find evidence for this in empirical psychology. (We mentioned the issue back in §1.5.) Empirical evidence for the quasi-spatial features of pictorial mental imagery has been gathered since the 1970s, including experimental work on mentally scanning images (Block 1983, Pinker 1980, Shephard and Metzler 1971). Such work showed, for instance, that the time taken to scan between two points of a mental image is often proportional to their subjective distance in the pictorial representation; that larger objects 'fill' the space sooner than smaller ones; and that the level of detail of the **(p.151)** depicted situations decreases at the periphery of the image, similarly to what happens with our visual field.

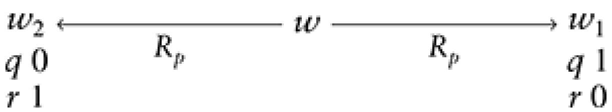
The view is disputed, however. Zenon Pylyshyn (1973, 1981, 2002) accepts the existence of mental images, but claims that their representational features can be reduced to non-pictorial, linguistically encoded representations. This 'imagery debate' is one of the most intractable controversies in contemporary cognitive psychology. If it resolves in favour of the quasi-spatial features of pictorial mental imagery, then we should include both SIMPLIFICATION and ADJUNCTION in the semantics. If Pylyshyn is right, and mental imagery can be reduced to linguistic representation, then we may want to drop SIMPLIFICATION or ADJUNCTION (or both).

7.4 The Under-Determinacy of Imagination

We move on to issues concerning disjunction. Imagination generally under-determines its contents: we imagine things vaguely, without this entailing that we imagine vague things. When you imagine Sherlock Holmes, by default you imagine him either left-handed or right-handed (or ambidextrous); but it's perfectly possible to not imagine him left-handed and not imagine him right handed either. Our semantics captures this:

$$[A](B \vee C) \nvDash [A]B \vee [A]C$$

For a countermodel, take three normal worlds w, w_1, w_2:



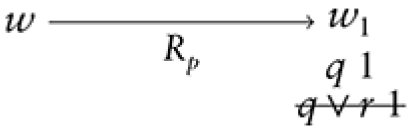
Every world R_p-accessible from w verifies $q \vee r$ and so w verifies $[p](q \vee r)$. But since w_1 doesn't verify r and and w_2 doesn't verify q, w verifies neither $[p]q$ nor $[p]r$.

Notice that this countermodel does not require non-prime worlds, which make a disjunction true without making either disjunct true. All **(p.152)** worlds in the countermodel are normal. (This isn't required, however: only w must be a normal world for the countermodel) The under-determinacy of imagination is delivered by the plurality of worlds accessed via R_A. Each accessible world may fill in the unspecified details in different ways. Since each $[A]$ is a universal modal operator, differences between the accessible worlds translate into unspecific acts of imagination.

We need worlds where disjunction behaves non-normally for another reason, however. When one imagines that B, one need not thereby imagine the disjunction of B and any arbitrary C (§7.1). One may lack the concepts involved in C. Our semantics captures this feature, too:

$$[A]B \nvDash [A](B \vee C)$$

For a countermodel, we need a normal world w and non-normal world w_1:



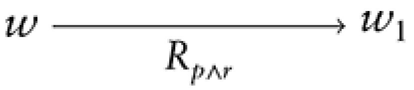
(Here, we've written ~~$q \vee r$ 1~~ to show that w_1 doesn't verify $q \vee r$. This is possible because w_1 is a non-normal world: ρ_{w_1} can relate any formula to any value or none.) In the model, by (S1[A]), w verifies $[p]q$ but not $[p](q \vee r)$. If we use our original clause (S2[A]) for falsity, then $[p]q \vee r$ is neither true nor false at w. But if we use the revised clause (S2[A]′), then $[p]q \vee r$ is false at w. Either way, the countermodel requires w_1 to be non-normal, else q's being true would force $q \vee r$ to be true.

7.5 Non-Monotonicity and Relevance

Imagination operators are non-monotonic, in the following sense:

$$[A]B \nvDash [A \wedge C]B$$

(p.153) To see why, consider this counterexample, with w a normal world:



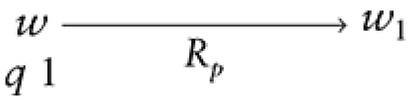
Since w_1 does not verify q, w does not verify $[p \wedge r]q$. But since no world is R_p-accessible, w trivially verifies $[p]A$ for every A, including $[p]q$. So $[p]q \nvDash [p \wedge r]q$.

As an act of imagination (in a given context) is individuated by its explicit content, one cannot in general import information into the explicit content itself without turning it into a different act. As you imagine Holmes walking across London, you imagine him walking across a British city. But if you imagine Holmes walking across London and that London has been displaced to France, you will not imagine him walking across a British city.

Other invalidities highlight the hyperintensional nature of imagination:

$$A ⥽ B \nvDash [A]B$$

Here's the counterexample, with w a normal and w_1 a non-normal world:



At every normal world (just w), if p then q, and hence w verifies $p ⥽ q$. (Recall that the strict conditional ⥽ looks at normal worlds only.) But this does not guarantee that, in imagining that p, we must thereby imagine that q. What we imagine may be impossible, as it is at w_1.

In particular, strict irrelevant conditionals fail the Variable Sharing Property from §6.1. But those conditionals do not entail the corresponding irrelevant imaginings. Here's an irrelevant but valid strict conditional:

$$\vDash (A \wedge \neg A) ⥽ B$$

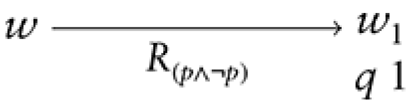
(p.154) But the corresponding imagination formula isn't valid:

$$\nvDash [A \wedge \neg A]B$$

and so

$$(A \wedge \neg A) ⥽ B \nvDash [A \wedge \neg A]B$$

Here's the counterexample, with w a normal and w_1 a non-normal world:



Imagining an inconsistent scenario does not trivialize our act of imagination. We can discriminate between different logical necessities and we do not imagine them all automatically whenever we explicitly imagine something. Thanks to non-normal worlds we have:

$$\Box B \nvDash [A]B \qquad \neg \Diamond A \nvDash [A]B$$

A constraint that should *not* hold in our semantics is (the counterpart of) what Lewis (1973b) called ‘Weak Centring’:

(WEAK CENTRING) If $w \in |A|$, then $w \in f_A w$.

This entails that, if a world w realizes the explicit content of an act of imagination A, then w is one of the worlds in the set outputted by the selection function for A. Even restricted to normal worlds, WEAK CENTRING validates a sort of modus ponens for imagination:

$$A, [A]B \models_? B$$

According to this principle, if the explicit content A of an act of imagination is true, and it is imagined in that act that B, then B also true. But this is wrong. Franz and Mark imagine that, in setting themselves a writing deadline, frenetic and productive writing will ensue ($[A]B$). They set the deadline (A); but things go on much as before for them ($\neg B$). Well-intentioned Brenda imagines, in her **(p. 155)** country leaving the EU, things getting better for her ($[A]B$). Her county does leave (A), but things get worse for Brenda ($\neg B$). So we should not accept WEAK CENTRING.

One way WEAK CENTRING can go wrong is when an agent imports false beliefs into her imaginings. We have spoken of importing *information* into imaginings, but we shouldn’t assume that what’s imported must be true. (If you think information is by definition factive, then we’re speaking of non-factive quasi-information.) ‘What people do not change when they create a counterfactual alternative [in imagination] depends on their beliefs’ (Byrne 2005, 10), and false beliefs may sneak in. You imagine Merkel signing treaties in Brussels ($[A]$), but you mistakenly believe Brussels to be in France. You import that belief and imagine Merkel to be signing treaties in France ($[A]B$). Merkel does in fact sign in Brussels (A), but this doesn’t imply Brussels is in France. In general, the role of the world w where the act of imagination takes place is to fix the agent’s beliefs, rather than to fix what is in fact the case.

7.6 Imaginative Equivalents

So far our logic of imagination is relatively weak, due to its highly hyperintensional features and the variability in the selection of the accessible worlds. We may add a *Principle of Imaginative Equivalents* (mimicking an analogous principle that holds in conditional logics (Priest 2008, 92)), whose effect is to limit the hyperintensionality of imagination:

(PIE) If $f_A \subseteq |B|$ and $f_B \subseteq |A|$, then $f_A w = f_B w$.

This says: if all the selected A-worlds make B true and vice versa, then A and B are ‘imaginative equivalents’. When we set out to imagine either, we look at the same set of worlds. Given (PIE), we have:

$$[A]B, [B]A, [A]C \models [B]C$$

(p.156) To see why, suppose normal world w verifies $[A]B$, $[B]A$, $[A]C$. By (S1[A]), we have $f_A w \subseteq |B|$, $f_B w \subseteq |A|$, and $f_A w \subseteq |C|$. (PIE) then gives us $f_A w = f_B w$ and hence $f_B w \subseteq |C|$. So by (S1[A]), w verifies $[B]C$.

This inference tells us that 'imaginative equivalents' A and B can be replaced *salva veritate* as modal indexes in [·]. Given the number of hyperintensional distinctions we can make in our imagination, there may be few imaginative equivalents for a given agent. But suppose that *bachelor* and *unmarried man* are imaginative equivalents for you (as they should be, we guess, for any competent English speaker): you are so firmly aware of their meaning the same that you cannot imagine someone being and not the other. So when you imagine that John is unmarried, you imagine that John is a bachelor ($[A]B$) and vice versa ($[B]A$). Suppose, in imagining that John is a bachelor, you imagine that he has no marriage allowance ($[B]C$). Then the same happens when you imagine that he is unmarried, $[A]C$.

(PIE) also licenses an inference we'll call *Special Transitivity*:

(ST) $[A]B$, $[A \wedge B]C \models [A]C$

To see why, suppose normal world w verifies $[A]B$ and $[A \wedge B]C$. Given OBTAINING, w also verifies $[A]A$ and so, by ADJUNCTION, $[A](A \wedge B)$. From OBTAINING and SIMPLIFICATION, we have $[A \wedge B](A \wedge B)$ and hence $[A \wedge B]A$ at all worlds. Then $f_A w \subseteq |A \wedge B|$ and $f_{A \wedge B} w \subseteq |A|$ and so, by (PIE), $f_A w = f_{A \wedge B} w$. Since w also verifies $[A \wedge B]C$, we have $f_{A \wedge B} \subseteq |C|$, hence $f_A w \subseteq |C|$, and so w verifies $[A]C$ too.

Should we accept (ST)? Some instances seem good. You imagine, in winning the lottery, having a lot of money ($[A]B$). You then imagine, in winning the lottery and having lots of money, that you'll have to pay a substantial amount of tax ($[A \wedge B]C$). It seems OK to infer *in the same context* that, in imagining winning the lottery, you thereby imagine having to pay a substantial amount of tax ($[A]C$). Of course, you can also imagine winning the lottery and avoiding paying any taxes. But that seems to create a different context from the one in which you imagined having to pay the taxes, on winning the lottery and so having lots of money.

(p.157) It may be that there are intuitive counterexamples to (ST) forceful enough for us to reject it. If so, we must drop either (PIE) or one (or both) of SIMPLIFICATION and ADJUNCTION. Perhaps what we should take from a counterexample to (ST), should one exist, is that imagination does not have a nature which respects conjunction after all (§7.3). If that's right, then we should drop either SIMPLIFICATION or ADJUNCTION (or both), without which the proof of (ST) will not go through, and retain (PIE) for the additional inferential power it gives us.

Chapter Summary

Imagination seems to have a logic, albeit one which is hyperintensional and sensitive to context (§7.1). We offered a semantics of imagination, with operators expressing ‘imaginative acts’ of mental simulation (§7.2). We then discussed a number of conditions we could impose on the semantics, in order to validate certain inferences (§7.3). One important issue is how acts of imagination interact with disjunction. One can imagine some disjunction as obtaining without being imaginatively specific about which disjunction obtains, for example (§7.4). We then turned to the issue of non-monotonicity: how B may follow from imagining that A, but not from imagining that $A \wedge C$ (§7.5). Finally, we discussed the *Principle of Imaginative Equivalents*, which, if valid, adds considerable power to the logic (§7.6). **(p.158)**