

POSSIBLE WORLDS SEMANTICS

Daniel Nolan. Draft. The final version of this chapter is in Russell, G. and Fara, D.G. (eds). 2012. *The Routledge Companion to the Philosophy of Language*. Routledge, London. pp 242-252. Please cite the final version.

One approach to specifying the meaning of pieces of languages is to treat those meanings as constructions out of possible worlds and possible objects. This technique is useful both in logic and in providing the semantics of natural languages. After introducing possible worlds semantics, this article will outline some of the applications that have convinced many philosophers and linguists of the usefulness of this framework, and will conclude with a discussion of one of the most conspicuous limitations of the framework, and some ways this limitation has been dealt with.

Extension and Intension

Pieces of language have extensions: predicates are associated with sets of objects that satisfy the predicate, sentences have truth-values, and singular referring expressions designate objects. Some aspects of the meanings of expressions seem to rely on more than the extensions of those expressions: extensions are often contrasted with “intensions.” In the context of possible worlds semantics, expressions are assigned possible-worlds intensions that are constructed from possible worlds and possible objects using set theoretic means. For example, the standard possible-worlds intension of a predicate is a function from possible worlds to extensions, that yields the predicate’s extension in each possible world. (These functions are sometimes called the *properties* associated with predicates.) One standard way of defining the intensions of declarative sentences is as a function from possible worlds to truth-values (a function that assigns a sentence True when it is true according to a world and False otherwise), or, as a simpler construction that carries the same information, the set of possible worlds where the sentence is true. These sets of possible worlds are often referred to as possible-worlds *propositions* expressed by sentences.

Possible Worlds Semantics in Logic and Language

The expression “possible worlds semantics” was first used to describe “semantics” in the logician’s sense. In this sense, possible worlds semantics is a matter of associating with a given logic a *model* that contains worlds; and assignments, relative to those worlds, of the truth-values of sentences, extensions of predicates, and so on. The best known application of possible worlds semantics is in the semantics of modal logics, usually attributed to Kripke 1959a and 1963, though see Copeland 2000 for the “pre-history” of this semantics.

Consider, for example, the propositional modal logic S4, which extends propositional logic with two operators applied to well-formed formulas: $\Box(\phi)$ for “it is necessary that ϕ ” and $\Diamond(\phi)$ for “it is possible that ϕ .” A model of S4 can be given by a ordered triple $\langle W, R, \nu \rangle$, where W is a set of worlds, R is a two-place relation on worlds, and ν is an evaluation function from propositional letters and worlds to the truth-values 1 and 0 (for True and False, respectively). R is an “accessibility relation” between worlds, and to ensure the model will model S4, we insist that R be both reflexive (so that for all $w \in W$, Rww) and transitive (so that for all $w, x, y \in W$, when Rwx and Rxy then Rwy). To specify ν , we first insist that ν assign exactly one of 1 and 0 to each propositional letter for each $w \in W$. We will represent ν assigning a propositional letter p the truth value 1 at world w as follows: $\nu_w(p)=1$, and in general assignments of truth-values at worlds will be represented in the form $\nu_\alpha(\phi)=\gamma$ for a world α , proposition ϕ and truth-value γ .

Next, we specify that ν assigns propositional-logical compounds truth-values, given the truth value assignments to propositional letters in each world, as those truth-values are built up in propositional logic. So, for example, when $\nu_w(p)=1$ and $\nu_w(q)=0$, then $\nu_w(p \ \& \ q)=0$, $\nu_w(p \ \vee \ q)=1$, $\nu_w(\sim p)=0$, $\nu_w(q \ \supset \ p)=0$, and so on. Finally, we specify how formulas with \Box and \Diamond are to be treated by ν . When ϕ is a well-formed formula, $\Box(\phi)$ is to be assigned the truth value 1 at a world w by ν ($\nu_w(\Box \phi)=1$) provided that all the worlds v such that Rwv are also such that $\nu_v(\phi)=1$. In other words, a sentence ϕ is specified to be

necessarily true at a world w if and only if in every possible world *accessible* from w (or “possible from” w) is a world where ϕ is true. The condition for \diamond is that $(\nu_w(\diamond\phi)=1)$ provided that *some* world v such that Rwv is also such that $\nu_v(\phi)=1$. In other words, ϕ is possibly true in a world w just in case some world “possible from” w is one where ϕ is true.

Propositional S4 is provably *sound* and *complete* with respect to these models. Whenever a conclusion Γ is provable from a set of premises Σ in S4, then every model which satisfies the above conditions and in which all the premises Σ are true at a world is one where Γ is also true at that world. Conversely, whenever it is that case that, in every model of the above sort, Γ is true at every world where all of Σ are true, then Γ is provable from Σ in the logic S4. Because of these facts, we can use the model theory to tell us which arguments are valid in S4 and which are not: if we can provide an S4 countermodel of an argument from some premises Σ to a conclusion Γ , then we can tell that Γ will not be provable from Σ in our S4 proof theory, for example.

Such models can be complicated to handle predicate calculus versions of S4, as well as still richer versions of S4. The model given above can also be varied to yield classes of models for other modal logics: in particular, we get quite a range of different modal logics just by putting different constraints on the R relation. Possible worlds models like these, whether for propositional logic or for predicate logic, have enabled modal logicians to see connections between a number of different systems, and in particular seem to have helped logicians to get a better understanding of the behaviour of *iterated* modal operators, such as in the formula $\Box\diamond\Box p$.

Models of this sort have been used in a number of different areas in logic besides the treatment of “necessarily” and “possibly”. Similar semantics have been offered for temporal logics (e.g. the logics of “until now”, “from now on” etc.), deontic logics (the logics of “ought”, “permitted”, and so on), and some epistemic and doxastic logics (logics of “it is known that...” or “... is (epistemically) justified”). Note that while the

formal structures are similar, the “worlds” in these models are often better thought of as times (in the temporal case) or epistemic states (in some epistemic logics). There is an extended sense of “possible worlds semantics” where the models for a logic are a possible worlds semantics provide they are similar, in the right kinds of ways, to the Kripke models for the logics of necessity and possibility. Compare the way that different mathematical systems are known as “geometries” even if they bear little relation to the structure of space, provided they bear certain kinds of similarities to systems like Euclid’s geometry.

Other kinds of relations on worlds have been found useful by philosophical logicians besides the accessibility relations of normal modal logics. Many non-classical Relevant Logics use a ternary relation on worlds to model the implication relations in those logics (see Routley et al 1983). “Conditional logics” like that of Stalnaker 1968 and Lewis 1973 use a “closeness” relation on worlds to model conditionals, particularly counterfactual conditionals. Interpretations of the language of probability sometimes employ *measures* on sets of worlds: this is one way to apply Kolmogorov’s mathematical theory of probability.

One thing that it is worth noting about the models used by logicians to investigate the formal properties of logics is that these models can be very useful even if they have nothing in particular to do with the project of providing the *meanings* of the logical operators and connectives: we could still employ the model theory to help with finding counterexamples to invalid inferences, for example. On the other hand, in linguistics and philosophy of language “semantics” is the study of the meaning of expressions (perhaps particularly that aspect of meaning which is in common across different uses of the same word). A number of linguists and philosophers have found it useful to investigate the semantics of natural languages, such as English, employing a framework of possible worlds. “Possible worlds semantics”, in this sense, involves using theories that postulate possible worlds in order either to give the meanings of e.g. English expressions, or at least to illuminate how those meanings work.

The two senses of “possible worlds semantics” are connected. Richard Montague introduced possible worlds into the systematic study of the semantics of natural language, and his approach was to take a simplified fragment of English, and then treat it as logicians had treated their artificial formal languages. Montague offered a model theory for his fragment of English that employed possible worlds, possible objects, and various set-theoretic constructions from them, in order to systematically yield assignments of truth-conditions to sentences: that is, assignments of truth-values to sentences relative to each world of the model. His reason for offering this model theory was to exhibit the way that the meanings of different English words go together to produce truth-conditions of English sentences. His semantics, in the model-theory sense, was the core of his semantics, in the sense of a theory about the meaning of English constructions. Montague’s system can be, and has been, applied to many other languages besides English, and Montague grammar remains a flourishing research program in linguistics. Montague 1974 is the best place to find Montague’s own contributions.

What Are Possible Worlds?

Possible worlds semantics relies on there being a domain of possible worlds, and usually things in those worlds to be the members of the sets associated with predicates in each world. So we have possible worlds and possible individuals. Does that mean that, in order to use possible worlds semantics, we need to think that there is an infinite range of alternate universes, full of merely possible individuals, including strange individuals like golden mountains and talking donkeys?

Some have argued that this is indeed the best way to understand possible worlds and possible individuals (Lewis 1986). More common is the view that possible worlds are some kind of abstract object. Whatever they are, they need to be able to have sentences or propositions “true according to” them, but there are a number of ways this might happen. A world literally containing a singing cow would have “A cow sings” true according to it in one straightforward way. But if worlds are representations they might be able to have “A cow sings” true according to them in the same way that a newspaper that contains the

sentence “A cow sings” can. And there are other options: perhaps a world composed of states of affairs that do not obtain could have as a component the state of affairs of a cow singing, for example. Or we could offer a modal analysis of “true according to”: perhaps a claim is true according to a world if, and only if, *were that world actualised*, the claim would be true. Finally, a theorist might resist analysis of the “true according to” relation altogether, and just take it to be an unanalysed fact that “A cow sings” is true according to one possible world but not according to another.

Some logicians and philosophers of language will be happy to leave the question of what possible worlds are, and what truth according to a world is, to the metaphysicians. But it is worth noting that the answer given to what possible worlds are, and in particular what the relation of truth-according-to-a-world is, may put constraints on what theoretical purposes possible worlds are suitable for. If the modal analysis of “true according to” is relied upon, then it will be much harder to treat the equivalence of “necessarily ϕ ” and “at all (accessible) possible worlds, ϕ ” as *explaining* necessity: we can use modality to explain what possible worlds are, or possible worlds to explain modality, but using each to explain the other risks circularity.

More worryingly, if we treat possible worlds as representations—as sets of sentences, or sets of propositions, or “maximal” propositions that settle every question—then we risk undermining their explanatory role in a theory of meaning. If we are explaining the meaning of sentences by associating them with sets of possible worlds, then on pain of circularity we should not take possible worlds to be sets of interpreted sentences, for example.

There are many other questions that could be raised about possible worlds, and indeed about possible objects. But for some purposes the metaphysics does not matter. For example, when using possible worlds semantics in the logicians sense, whether the members of W or the members of the domains associated with each $w \in W$ are anything like real possible worlds and the full range of merely possible objects is irrelevant from the technical point of view.

Indeed, even if we are using possible worlds for the project of semantics in the linguistics sense, we may be able to treat them as only formal modeling devices. Whether or not we can do so will depend on what sort of explanations they are meant to provide—if they are invoked only for modeling structural constraints on meanings, it may not matter what the objects in the models are. On the other hand, if we take ourselves to be illuminating the meaning of expressions by providing genuine truth-conditions for sentences, then the “worlds” and constructions from them should arguably have something to do with what can possibly be the case.

Possible Worlds Semantics for Natural Language: A Simple Model

Each kind of linguistic construction is assigned an “intension” or “semantic value” constructed from possible worlds, possible objects, and sets ultimately built up from these. *Sentences* are assigned, as their intensions, sets of possible worlds: these are the worlds at which the sentences are true, so in assigning these sets we are specifying the conditions under which the sentences are true. These possible-world truth conditions are identified as *propositions* in this framework. *Predicates* are assigned functions from worlds to sets of possible objects: the set associated with a predicate and a world is the extension of that predicate at that world. These functions from worlds to extensions are identified as *properties* in this framework. For an object to *have* a property, in a world, is for the object to be in the extension of that property at that world.

We can provide intensions for quantifiers: the intension of the universal and existential quantifiers are certain function from worlds to sets of *properties*, in the above sense. The universal quantifier, for example, is assigned, at each world, the set of properties in common to *all* the objects at that world, and the existential quantifier is assigned the set of properties such that *at least one* thing at that world has that property.

The intensions of larger linguistic structures are functions of the intensions of their components, so e.g. we can build up the intension of a sentence from the intensions of its constituents. The rule for a sentence of the form [quantifier];[predicate], for example, is

that a world w is in the intension of the sentence if the intension of the quantifier is a function from w to a set of properties that includes the property associated with the predicate. This may become clearer with an example: suppose our sentence is “something is hungry”. When we evaluate a world, say w_5 , the intension of “something” will associate with w_5 a set of properties (the set of properties had by at least one thing in w_5). Suppose the property that is the intension of “is hungry” is a member of that set. Then w_5 will be one of the members of the proposition associated with “something is hungry”. Which is as it should be, since those conditions ensure that one of the objects in w_5 is in the extension of the property associated with “is hungry”. If one of the objects has the property of being hungry in a world, we want “something is hungry” to count as true there.

If we wanted uniform principles about how different semantic categories went together, and the two kinds of basic sentences we allowed were [quantifier];[predicate] and [name];[predicate] sentences, then it would make sense to give names the same general kind of semantic values as quantifiers. Montague 1973 did this, treating names as also being associated with sets of properties: “Barack Obama”, for example, would be associated with a function from worlds to the set of properties that Barack Obama has at those worlds. “Barack Obama is hungry” would then have a world w in its intension just in case the set of properties associated with “Barack Obama” at w included the property associated with “is hungry”. Alternatively, we could have different rules for different kinds of sentences: e.g. we could allow the semantic value of “Barack Obama” just to be an individual, and the rule for [name];[predicate] sentences to be that a world w is in the intension associated with such a sentence provided that the individual that is the intension of the name is in the extension, at w , of the property associated with the predicate.

Extensional connectives such as “and”, “or” and “not” can be assigned intensions: “and”, for example, can take two propositions as inputs and yield a proposition as output (it, in effect, functions as set-theoretic intersection on sets of worlds: ϕ and φ is true at the set of worlds that is the intersection of the set of worlds where ϕ is true and the one where φ is true). Intensional operators like the sentential adverbs “Necessarily” and “Possibly”

can also be treated as functions from propositions to propositions. The simplest rules are these (where ϕ is a sentence): *Necessarily* ϕ has, as its intension, the proposition containing every world just in case the intension of ϕ includes every world, and has a null intension otherwise. *Possibly* ϕ has, as its intension, the proposition containing every world just in case the intension of ϕ includes at least one world, and has the null set as its intension otherwise. More complicated rules involving accessibility relations need to be invoked if more complex notions of necessity and possibility are in play.

Obviously the simple theory presented so far barely scratches the surface of the complexity of natural language: it does not even handle tense or the sorts of sentences we would represent in predicate calculus with multiple variables. The interaction of the syntax and semantics has not been addressed either. But hopefully it gives the flavour of how such a theory might be set up. To see how a possible worlds semantics might earn its keep, we should discuss some more difficult constructions that possible worlds semantics has had some success with.

Some Applications of Possible Worlds Semantics for Natural Language

Possible worlds semantics is useful, as we have seen above, for handling sentential adverbs like “necessarily” and “possibly”. It is useful not just for an “unrestricted” sense of necessity and possibility, corresponding to truth in every possible world or some possible world, respectively. It is also useful when dealing with a restricted necessity, such as saying what is necessary given the laws of nature, or what is feasible given some other constraints. (This corresponds to a restriction on the set of worlds relevant for the semantic value.) As was noted in the section above on logic, possible worlds semantics has also been found very useful in dealing with the so-called “deontic modalities” such as “It is obligatory that...” or “It is permitted that...” and their more natural-sounding English relatives, and also for dealing with so-called epistemic modalities used for specifying what is known or what is justifiably believed, or what “has to be” or “might be” in the epistemic sense of those expressions. (See Kratzer 1981 for a possible-worlds treatment of “ought”.)

Possible worlds semantics come into their own for a range of non-extensional uses of language. For example, consider so-called “intensional transitive verbs”. Examples include “seeks”, “owes”, “wants”, “offers”, and a number of others. Whether “John seeks a city of gold” is true or not does not just depend on the extensions of “John” and “a city of gold”. Suppose, as is plausible, that there are no cities of gold, and no flying horses. “a city of gold” and “a flying horse” would thus plausibly have the same extension: but “John seeks a city of gold” can be true while “John seeks a flying horse” is false, so “seeks” is intensional. If instead of extensions we use possible world intensions, we can draw the distinction we need. There are possible cities of gold, and possible flying horses, so “a city of gold” and “a flying horse” have different intensions—you could see those intensions as being *the property of being a city of gold* and *the property of being a flying horse*, respectively. Since the intensions are different, if we treat the intension of “seeks” as, in this case, a function from the two intensions that flank it to a proposition, then we can get the right answer, in this example, that the first sentence has the truth-value True at the actual world and the second does not.

One of the applications that caught the attention of the philosophical community following Stalnaker 1968 and Lewis 1973 was the use of possible worlds semantics to give a theory of the truth-conditions for conditional sentences. According to Stalnaker’s proposal, for example, “if A then B” is true at a world w just in case a “selection function”, which is a function from propositions and worlds to worlds, when it takes as arguments the proposition associated with A and the world w , yields a world x that is a member of the proposition associated with B. The effect of this is that the conditional is true at a world if the “closest” world to w where A is true is also one where B is true. This formal proposal yields a variety of interesting logics of conditionals depending on what constraints are put on the selection function, and deals with a number of problems that faced traditional theories of conditionals. For example, for “if A then B” to be true seems to require more than that either A is false or B is true (the “material conditional” account): I am currently sitting down, but it does not seem true that “if I had stood up, I would have exploded”. On the other hand, requiring that for “if A then B” to be true B has to be true in *every* possible world where A is true is too strong a requirement: “if I

had stood up, I would *not* have exploded” seems true, though there is nothing absolutely impossible about a situation where I stand up and then explode. A lot of contemporary work in exploring the semantics of conditionals employs some relative of this approach.

A third important application of possible worlds semantics is in the semantics of *propositional attitude* ascriptions. Sentences like “John believes the number of planets is 10” or “Mary desires that everyone love someone” seem to require a relationship between a person and a proposition: the propositions expressed by “the number of planets is 10” and “everyone love someone”, respectively. Constructions like “believes that” are not extensional: substituting one true sentence for another in the scope of “believes that” does not, in general, preserve truth-value. Furthermore, the behaviour of propositional attitude contexts is too complex to just treat “that ϕ ” constructions as, in effect, referring to propositions. “There is someone such that everyone believes that he is the murderer” cannot be treated as referring to a proposition with “that he is the murderer”, since “he” is in effect bound by “There is someone”. Moving to a possible worlds semantics allows phenomena like this to be dealt with better than existing extensional frameworks.

The above only scratches the surface of work done in the possible worlds tradition. Using possible worlds semantics arguably illuminates topics in semantics as diverse as adverbs, progressives, dispositional expressions, the semantics of embedded questions, and plays a vital role in theorising about effect of context, including in the development of dynamic semantics. For an introductory discussion to more applications than could be discussed here, see Partee 1989, especially section III.

The Problem of Necessary Co-Extension

Some sentences are not true at any possible worlds, and some predicates are true of any possible objects (“is round and not-round” might serve as an example). Furthermore, some pairs of sentences are true in some possible worlds, but are true in all the *same* possible worlds, and some pairs of predicates hold of exactly the *same* possible objects. Since orthodox possible worlds semantics takes sentences true at the same worlds to have the same intensions, and predicates true of the same possible objects to share intensions,

this leads to some surprising results.

The possible worlds semantics treatment of intensional transitive verbs, for example, misfires when impossibilities get involved. Or at least the straightforward version of possible worlds semantics does. Suppose Alan is seeking a proof of the rationality of π . Such a proof is impossible, so the set of objects, across possible worlds, that meet the condition of being a proof of the rationality of π is the null set. Now, suppose Ben is seeking a round square: again, given the impossibility of round squares, the condition of being a round square corresponds to the null set. “proof of the rationality of π ” and “round square” have the same possible-worlds intension. So we should be able to conclude, by substitution of equivalences, that Alan is seeking a round square, and Ben is seeking a proof of the rationality of π . But that substitution is not valid: Alan seems well within his rights to deny that he is seeking a round square, his behaviour is not that of someone who is seeking a round square, and our psychological insight has failed if we report that Alan is doing that. Similar problems arise for verb phrases such as “prays to”, “try to”, “wish for”, “offer”, and even “owe”—I might owe you a proof of the rationality of π , if I promised to deliver one.

Propositional attitude ascriptions also are dealt with badly by the basic possible worlds semantic treatment: anyone who believes one strictly impossible thing (thing true in no possible world) believes them all, for example. As well as the problem of impossible contents, propositional attitude ascriptions have problems with necessarily true contents: anyone who believes one necessarily true thing believes them all, according to this model, but even if mathematical truths are necessary, they are not that easy to discover! There are also problems with beliefs about contingent matters where what seem to be two different belief contents obtain in the same possible worlds. Suppose Robin Hood and Robin of Locksley are the same person, though few suspect this. The Sherriff of Nottingham can believe that Robin Hood robbed the abbot, while not believing that Robin of Locksley robbed the abbot. Or so we would ordinarily think. If “Robin Hood” and “Robin of Locksley” are *rigid designators*, and pick out the same individual across possible worlds, then the two names have the same possible worlds intension. So,

according to the straightforward possible worlds semantics for names at least, “The Sherriff believes Robin Hood robbed an abbot” is necessarily equivalent to “The Sherriff believes Robin of Locksley robbed an abbot”: the Sherriff cannot believe one without believing the other. A similar problem will arise for natural kind terms, given the usual theory of how they work—anyone who believes there is water in their glass also believes there is H₂O in their glass, according to a straightforward possible worlds semantics. But not everyone believes that water is H₂O.

Finally, the orthodox possible worlds treatment of conditional statements gives results that many find odd. Conditionals with necessarily false antecedents that share a consequent must all be treated the same, since those antecedents all share an intension: the standard line (e.g. Lewis 1973 pp 24-26) is to treat all conditionals with necessarily false antecedents as true at all worlds, regardless of the consequent. However, many people give different verdicts on different “counterpossible” conditionals: “If 8 had been divisible without remainder only by 1 and itself, it would have been prime” seems good, but “If 8 had been divisible without remainder only by 1, 2, 5 and itself, it would have been prime” looks less appealing: in the latter case it is tempting to say it would not have been prime, but it would have been composite instead. (See Nolan 1997 for a case for positing counterpossible conditionals with non-trivial truth values.)

There are a number of responses available to advocates of possible worlds semantics here. The most straightforward is to bite the bullet: Alan *is* seeking a round square, though he might not put it that way; the Sherriff *does* believe Robin of Locksley robbed an abbot; everyone always knew that water was H₂O, and so on. Biting these bullets often goes with a strategy of explaining the *pragmatics* of these utterances so that they are somehow misleading or incorrect to say, even if they are not false (this seems to be the proposal about some cases in Stalnaker 1978, for example).

Another relatively straightforward response is to say that possible worlds semantics is a *model* of the meanings of linguistic expressions that is useful for some purposes but not others: perhaps it is analogous to Newtonian mechanics, which is a good enough

approximation when it comes to bridge-building, but not the right physics to use when designing particle accelerators. If this approach is taken, it leaves open the question about what the more fully accurate theory might look like—but for some projects, it makes sense to put that question off.

Another response is to say that a semantics of natural language should use constructions from possible worlds as *one* of the components of a “meaning”: perhaps semantic values should be pairs of possible-worlds-constructions and something more linguistic, for example, or that possible worlds intensions should be components in a “structured meaning” that looks rather like a syntactic structure, as in Creswell 1985. There are potentially as many different versions of this strategy as there are components we could add to a semantic value, so it is difficult to generalise about the prospects of this way of developing possible worlds semantics.

A way of complicating the postulated intensions without bringing in things besides constructions from possible worlds is to treat intensions as “two dimensional intensions”. David Kaplan pioneered this move to deal with demonstratives and indexicals: expressions like “that” or “I” or “tomorrow” which make different contributions to truth-conditions on different occasions of use (Kaplan 1989). “It will rain tomorrow” can be true when I say it today but false when I say it tomorrow, for example. Kaplan held that the meaning of an expression like “that” or “tomorrow” had two aspects: a *content*, which was in effect a standard possible-worlds intension, and a *character*, which was a function from certain n-tuples (with components like the speaker, and the time of utterance) to contents. Such n-tuples can be generalised to *worlds* of utterance, or worlds of utterance plus a “centre” to indicate which speaker, time, audience etc. in the world is being picked out - these are sometimes called the ‘context’ of the utterance.

While Kaplan intended his framework for a limited range of context-dependent expressions, later “two-dimensional” theorists have offered similar treatments for proper names and natural kind terms. The *contents* of proper names pick out the same object in each world, like a standard possible-worlds intension, but the character associated with a

name maps the name to different contents depending on features like which object is at the source of the reference-chain associated with the indicated speaker, or the referential intentions of the designated speaker. In this framework, there are contexts which take “Robin Hood” and “Robin of Locksley” to different contents, and since the overall functions-from-contexts-to-content intensions are thus different for the two names, there is scope for those intensions to interact differently with pieces of language such as “believes that”. (See Garcia-Carpintero and Macia, 2006 for a collection containing a number of papers for and against such generalised two-dimensional treatments.)

A final kind of modification for orthodox possible worlds semantics to deal with the problem of necessary co-extension is to be more generous with what worlds the theory allows. We got the problems because some expressions had equivalent applications in every possible world: “a proof of the rationality of π ” and “a round square”, for example, or “Robin Hood” and “Robin of Locksley”. If we allowed worlds to have the extensions of these expressions come apart from each other, these problems would not arise.

One way to do this, which is tempting in the case of propositional attitude ascriptions and intensional transitives involving the activity of agents, is to use *epistemically* possible worlds or *doxastically* possible worlds. These worlds might be more generous than standard possible worlds: for example, if it is possible to believe ϕ without believing φ , there are “worlds” where ϕ and φ come apart, and connectedly that the extensions of two expressions (like “water” and “H₂O” or “proof of the rationality of π ” and “round square”) come apart at some world if it is possible to believe that an object is in one extension but not in the other. Or one could go even further and drop the restriction to do with our capacities to believe: if a worlds semantics allowed arbitrary impossible worlds as well as possible worlds, the problem of necessary co-extension could be guaranteed to not arise.

Possible worlds semantics has provided a powerful and impressive framework for systematically dealing with the semantics of puzzling constructions in language. Whether the theory of semantics it yields needs to be extended by other devices, and if so what

those devices are, will continue to be one of the important debates in the philosophy of language in the coming decades.

Related Topics: Intensions, Extensions, Characters and Beyond; Modal Logic; Montague Grammar; Names; Necessity; Propositional Attitudes.

References

Copeland, B.J. 2000. "The Genesis of Possible Worlds Semantics," *The Journal of Philosophical Logic* 31: 99-137

Cresswell, M. 1985. *Structured Meanings*. MIT Press, Cambridge MA.

Garcia-Carpintero, M. and Macia, J.(eds.) 2006. *Two-Dimensional Semantics: Foundations and Applications*. Oxford University Press, Oxford.

Kaplan, D. 1989. "Demonstratives," in Almog, J., Perry, J. and Wettstein, H. *Themes from Kaplan*. Oxford University Press, Oxford. pp. 481-563.

Kratzer, A. 1981. "The Notional Category of Modality," in Eikmeyer, H.J. and Rieser H. (eds.) *Words, Worlds, and Contexts*, de Gruyter, Berlin: 38-74.

Kripke, S. 1959a. "A Completeness Theorem in Modal Logic," *Journal of Symbolic Logic* 24: 1-14

Kripke, S. 1963. "Semantical Considerations on Modal Logic," *Acta Philosophica Fennica* 16: 83-94

Lewis, D. 1973. *Counterfactuals*. Blackwell, Oxford.

Lewis, D. 1986. *On the Plurality of Worlds*. Blackwell, Oxford.

Montague, R. 1973. "The Proper Treatment of Quantification in Ordinary English," reprinted in Montague 1974 pp 247-270.

Montague, R. 1974. *Formal Philosophy: Selected Papers of Richard Montague*. (Thomason, R. ed.) Yale University Press, New Haven.

Partee, B. 1989. "Possible Worlds in Model-Theoretic Semantics: A Linguistic Perspective," in Allen, S.(ed.) *Possible Worlds in Humanities, Arts and Sciences. Proceedings of Nobel Symposium 65*. Walter de Gruyter, New York. pp 93-123.

Routley, R., Meyer, R.K., Plumwood, V. and Brady, R. 1983. *Relevant Logics and its Rivals*, Volume I. Ridgeview, Atascadero, CA.

Stalnaker, R. 1968. "A Theory of Conditionals," in Rescher, N (ed.) *Studies in Logical Theory*, American Philosophical Quarterly Monograph Series, no.2. Basil Blackwell, Oxford. pp 98-112

Stalnaker, R. 1978. "Assertion," in Cole, P. (ed). *Syntax and Semantics: Pragmatics* Vol. 9. Academic Press, New York. pp 315-332