1 Introduction

Have you ever wondered why you use a safety belt when travelling by car? Or why a child decides against stealing something they desperately want yet their parents wouldn't buy? Modal cognition is typically at the core of such decisions. We know that any of us *could* have a car accident and that a properly worn safety belt decreases human harm in case there is one. This knowledge informs our decision to use the safety belt. (Or, for otherwise inclined subjects, the *possibility* of a fine would play a similar role.) Analogously, the child knows they *might* be caught, and the fear of punishment if they are caught has the power to bend their immature moral will. We usually *take it for granted* that these decisions are rational, and our doing so is partly explained by our belief that those possibilities *are* so and that modal cognisers are, by and large, justified in their modal beliefs.

But *what is modality* in the first place? This is the question at the core of the present Element, and I shall approach it in a two-tier manner, with cumulative effect: first, I focus on modality in general and, after that, on metaphysical modality in particular – arguably the kind of modality which is distinctively central to metaphysics.

Crudely put, modality refers to modes; modes, for instance, in which a proposition can be true or a property possessed: necessarily or contingently. More generally, modality in philosophy is the area that studies the notions of *necessity*, *possibility*, and *contingency* (plus cognates). But modality is a vast phenomenon and, within it, we can distinguish different types, even families, such as deontic, epistemic, and alethic. And depending on which of these families is our focus on a given occasion, our modal investigations will also lead us into studying related notions such as *essence*, *accident*, *disposition*, *power*, *permission*, *obligation*, *norm*, *knowledge*, and *evidence*, among others.

Which family of modalities is most central in a given philosophical context depends on the branch of philosophy under consideration: deontic modalities, for instance, are likely to be more central to ethics than to general epistemology. In metaphysics – the series in which this Element is situated – alethic modalities are largely central and, among them, *metaphysical* modality is distinctively so. But modality as a *general* phenomenon already raises very interesting questions, including metaphysical ones, that wouldn't vividly manifest if one were to focus on alethic or, within it, metaphysical modality from the outset. Section 2 of this Element is thus devoted to the general phenomenon. In it, we will look into questions such as what it is that brings all different kinds of modality together: what is it, if anything, that makes all kinds of modality, kinds of *modality*? Is there a fundamental modality among them? The contemporary

literature is interestingly divided about this and, by way of anticipation, we will see that positions on the matter are not independent of one's stance on the source of metaphysical modality. This section will thus inform Section 3, which focuses, more narrowly, on different views on the source of *metaphysical* modality. Section 3 will also evidence (albeit only in passing) how questions about the metaphysics of modality can be impacted by epistemological concerns; i.e., concerns about our knowledge of modality.

2 Modalities

2.1 Variety of Modal Discourses

Modal discourse is ubiquitous and it is very much entrenched in our daily lives. We have considered a couple of examples in the Introduction, and it is not difficult to come up with a much longer list of scenarios where modal propositions (or modal beliefs) are at the core of our reasoning or decision-making processes. As crudely characterised in the Introduction, a modal proposition involves some modal notion – central among them are *possibility* and *necessity*. In the following, it is the absence of modal notions in (1) and their presence in (2) that makes only the second proposition a modal one:[1]

(1) Iris' knee is bleeding.
(2) The fall *must* have been worse than it looked.

As with any other type of proposition, modal propositions can be true or false. Regardless of their truth-value, all the propositions below are modal in virtue of each involving a modal notion:

(3) Iris *could* be more careful.
(4) Two plus five *must* be seven.
(5) One *can't* always be a Good Samaritan.
(6) Being a philosopher is *accidental* to Socrates.
(7) A rocket *can* travel faster than light.
(8) *Necessarily*, if Ernest isn't married, he is a bachelor.
(9) You *can* drive without a safety belt when you're heavily pregnant.
(10) If Ernest isn't married, he's *necessarily* a bachelor.
(11) Umar *can't* be at the hospital.
(12) You *can* jump this stream.
(13) You *must* respect people's turns (e.g., in a queue).
(14) It *must* have been raining.
(15) You *can't* torture people.

---

[1] It will do no harm here not to distinguish between propositions and sentences expressing them.

These propositions, and some contrasts among them, will first serve us to illustrate the ubiquity of modal discourse, and they will then help us draw distinctions between different types of modalities (Sections 2.2, 2.3, and 2.4). Following this, we will tackle the much-contested issue of *unification* (Sections 2.5 and 2.6) and use this discussion to reflect on the contemporary role of *possible worlds* (Section 2.7) and the prospects of precisifying, within modality, the *de re/de dicto* distinction (Section 2.8).

Of the propositions above, some might seem artificial (they have been chosen with a purpose) but others are ordinary modal propositions and it's not difficult to conceive of scenarios where one might easily apply them. For example, someone tells you that your neighbour, Umar, was admitted into hospital yesterday night with severe Covid-19; this surprises you as you've just spoken to him on the phone this morning. You reply with (11) – Umar *can't* be at the hospital. In a second example, your daughter is playing in the park, which is divided by a stream. She's on one side, but wants to join a group of children who are playing on the other side. She's impatient, and wonders whether to walk down to the bridge, cross it, and walk up again, or whether she should dare jumping across the stream. After assessing the circumstances, you encourage her with (12) – You *can* jump this stream. And a third example: your nephew enters a busy candy shop and goes straight to the counter demanding a lollipop. You intervene with (13) – You *must* respect people's turns.

These are just a couple of examples. They may be trivial, but they are illustrative of how quotidian modal thought and modal reasoning is. Beyond quotidian life, preventive medicine is a health specialisation whose goal is to prevent people from contracting illnesses they *can* (relatively easily) contract. At a more abstract level, modal reasoning is central in scientific methodology. The process of acquiring information about how the world *is* can be seen as the process of ruling out that the world is some other *possible* way. Let us suppose that there are two hypotheses, H_1 and H_2, that seem to equally explain our body of evidence, E. For all we know, the world could be like H_1 says, but it also could be like H_2 says. If we want to know how the world is, and given that evidence E is neutral between H_1 and H_2, we need disambiguating evidence. This extra evidence can be achieved by designing and carrying out experiments; this is common scientific practice. And yet, designing and carrying out experiments can be very costly (financially and otherwise). Could we justify this practice if we had sufficient reasons to believe that either H_1 or H_2 are impossible? Or if we had sufficient reasons to rule one out? When a hypothesis is taken seriously, we are assuming, at least implicitly, that the world *could* be the way the hypothesis describes it.

4   *Metaphysics*

The reader who is already familiar with the distinction between *alethic*, *epistemic*, and *deontic* modalities might be wondering why I'm not being more rigorous. Is it epistemic or alethic possibilities that are being assumed in scientific practice, for instance? Is it in an epistemic sense or in a physical sense that Umar can't be at the hospital? However, this lack of rigour, at this stage, is intended. It illustrates that there are indeed different senses associated with our modal vocabulary, and which one is involved easily varies from context to context (including a speaker's intentions). Consider again proposition (11) – Umar *can't* be at the hospital. Is this true or false? This question cannot be answered once and for all; it depends on the sense (and context) in which it is uttered. For instance, above, I pictured a context intended to be one where an utterance of (11) is very likely to be true. *If you know* that you've heard Umar at home this morning, talking on the phone, then (it is likely that) *he couldn't* have been admitted into hospital with severe Covid-19 the evening prior. There *must* have been a misunderstanding somewhere: maybe he was checked but not kept in? The sense of modality in which these conclusions are true is, as we shall see shortly, an epistemic one: it is dependent on what we know.

But (11) could be uttered in other contexts where we don't mean an epistemic sense of 'can'. Imagine a group of friends and one of them, Shyla, asks where Umar is. Joel replies: 'Well, he's quite healthy, so he can't be in the hospital', to which Kiran protests that his being quite healthy doesn't mean that he cannot fall ill or have an accident. In this context, (11), as uttered by Joel, is likely to be false, and what makes it false is not dependent on what the group of friends *know* or *believe about* Umar, but rather on what's modally *true of* him: regardless of his degree of health, Umar, as an organism, is not immune to illnesses or accidents. As we will also see shortly, this dependence on *truth* makes this sense of modality an alethic one. In addition to the epistemic and the alethic, there's yet another kind of modality that we will also focus on in this section: the deontic one, which has to do with *norms*. To get a taste of it, consider (15) – You *can't* torture people. If this strikes you as true, you're most likely reading it in the deontic sense, and, if so, the message you're taking from it is that it's not morally permissible to torture people. Notice that, in an alethic sense, (15) is false: as too often witnessed, there aren't many physical impediments to people torturing others.

It is the aim of the next three sections (Sections 2.2, 2.3, and 2.4) to unfold these brief remarks by focusing on three main families of modalities: epistemic, deontic, and alethic. I shall first introduce them separately. After this, in Sections 2.5 and 2.6, I will be most interested in exploring the relations between them, paying special attention to debates in recent literature having to do with the prospects of finding a unifying treatment. Before all this, let me introduce the notions of *possibility* and *necessity* as duals of one another. Roughly, by

'duality' in this context we mean that we can define *possibility* in terms of *necessity*, and vice versa.[2] Specifically, *p is possible* – in symbols, ◊p – means that it is not necessary that not p – in symbols, ¬□¬p. We have just defined possibility in terms of necessity. And we can correspondingly define necessity in terms of possibility: *p is necessary* (□p) means that it is not possible that not p (¬◊¬p). With duality in place, let us turn to the different types of modalities.

2.2 Epistemic Modalities

Assuming the duality between *possibility* and *necessity*, it is open to us to directly characterise the *epistemic* sense of only one of the two modal notions – necessity or possibility – and to arrive at the other by means of the conceptual path that their duality affords. I shall illustrate this as we move along, starting with the notion of epistemic *possibility*. As a first approximation, epistemic possibility can be characterised in the following way:

Epistemic possibility (EP):
A proposition, *p*, is epistemically possible for a subject, S, if and only if *p* is (logically) compatible with all S knows.

It is this dependence on *knowledge* that gives epistemic modality its name, from the ancient Greek *episteme*, meaning knowledge or science. The first thing to note is that (EP) characterises *epistemic possibility* as relative to a subject. This cannot be otherwise due, precisely, to its dependence on a given subject's body of knowledge. Note also that, since different subjects will clearly know different things, which propositions are (or are not) epistemically possible will vary accordingly across subjects. To illustrate: if I have no idea what is the highest speed a spur-thighed tortoise can reach, it's epistemically possible *for me* that Clark (my friend's spur-thighed tortoise) walked at three kilometres per hour yesterday; such speed is certainly compatible with all I know – virtually nothing – about tortoises. However, a herpetologist will have relevant information about tortoises that puts her in a position to rule out that Clark ever reached that speed. Thus, it is *not* epistemically possible *for the herpetologist* that Clark walked at three kilometres per hour yesterday.

   (EP) is a good first approximation. It captures a kind of epistemic modality often made salient in ordinary contexts, as the Umar example above and the tortoise case just described, serve to illustrate. Given that we're assuming the duality of the notions of necessity and possibility, we now have two options to characterise the corresponding *epistemic necessity*: we either characterise it

---

[2] This is widely but not universally accepted. Arthur Prior, for instance, developed a modal logic where □p is not even *extensionally* equivalent to ¬◊¬p. (See system Q in Prior 1957.)

directly, or else we do it from (EP), exploiting the duality. The two ways will (and should, if the duality is such) be extensionally equivalent:

Epistemic Necessity (EN), directly:
A proposition, *p*, is epistemically necessary for S if and only if *p* (logically) follows from what S knows.

Epistemic Necessity (EN), exploiting the duality:
A proposition, *p*, is epistemically necessary for S if and only if it is not epistemically possible for S that not *p*.

(EP) and (EN) jointly characterise what is perhaps the most paradigmatic type of epistemic modality. To be rigorous, however, they characterise potentially as many different epistemic modalities as there are individuals (or, more rigorously, as there are pairs of individuals and times). For what we've got is a *schema*, rather than a modality, and we can replace 'S' by any individual, each of these moves resulting in an epistemic modality.

But even leaving this schematicity to one side, this schema (of epistemic modalities relativised to one single individual) is not the only one within the family of epistemic modalities, not even the one that will always be salient. Sometimes, the context might make salient instead an epistemic modality relativised to a *group* of individuals. To capture these other types of epistemic modalities, we would need to mention explicitly a group of individuals in the schema. In addition, we would need to accompany this by an unambiguous notion of collective knowledge. This gives rise to alternative, non-equivalent (schemas of) epistemic modalities, as the following illustrates without being exhaustive:

*Epistemic Possibility, relativised to groups, **intersection**:*
A proposition, p, is epistemically possible for a group of people *G* if and only if *p* is (logically) compatible with what every member in *G* knows.

*Epistemic Possibility, relativised to groups, **union**:*
A proposition, *p*, is epistemically possible for a group of people *G* if and only if *p* is (logically) compatible with what *some* member in *G* knows.

As above, these are schemas. In order to get an epistemic modality from them, we would need to plug in a given group of people in each case. Technically, for any group of people we can think of, we could plug it in, and we would obtain a specific epistemic modality as a result. Admittedly, some of these modalities are of greater interest than others. For instance, in the middle of a pandemic, plugging in a worldwide group of epidemiologists would generate a much more interesting epistemic modality (in either the union or the intersection sense) than the group of members in my household, especially with respect to possibilities

concerning, for instance, how stressed hospitals will be in a year's time, or how risky it is to book a flight for two months' time.

We have focused so far on epistemic modalities in a strict sense: that is, modalities that are dependent on *knowledge*. One might wonder what the result would be if we considered alternative notions in the vicinity, such as *belief*, or even *justified belief*. The answer, simply put, is that we would get different types of modalities which, with a suitably relaxed understanding of the label 'epistemic modality', might still be taken to be such, and for which, with a more stringent insistence on labels, 'doxastic modalities' would be more apt. Regardless of our terminological choices, all these modalities have in common their relativisation to an individual or group of individuals, and that what is possible or necessary for them depends on the *body of information* that the individuals are committed to, at least at the level of belief. This makes all the elements in this (relaxed) group of modalities quite close to one another as phenomena.

And yet, one interesting difference between epistemic modalities and doxastic ones (in stringent senses of these labels), stems from the facticity of knowledge and the lack of facticity of belief (even of justified belief). We know that what is known is indeed the case – knowledge is factive – but that beliefs, even justified ones, can be false – belief is not factive. This difference has a sharp impact on the *modal logics* that can be used to model epistemic and doxastic modalities. Think of the principle according to which *what's necessary is (also) the case*, which is known as principle (M) in modal logic (in symbols, $\Box p \rightarrow p$). With a *stringent-epistemic* sense of 'necessary' this principle comes out as *valid* (cannot fail to be true), but this is not the case with the *doxastic* sense of the term. We will take the cases in turn. If p is *epistemically* necessary (if $\Box p$, in the epistemic sense), then p follows from a body of knowledge (be that knowledge of a single individual or of a group of individuals). Since knowledge is factive, p is the case. As a result, if $\Box p$ then p. And since the choice of p was arbitrary, this concludes the reasoning for the epistemic sense. By contrast, if p is *doxastically* necessary (if $\Box p$, in the doxastic sense), then certainly p follows from a body of beliefs, but this is insufficient to guarantee the truth of p; indeed, p could be falsely believed (that is, the body of information from where p follows could be false). As a result, in the doxastic case, we cannot safely transition from $\Box p$ to p: some of these transitions could lead us from truth to falsehood. Things being so, if one is to model *epistemic* modality with modal logic, it will need to be with a modal logic that includes principle (M), whereas one would need to exclude it if one is modelling *doxastic* modalities.[3] We will return to (M) when considering deontic modalities, to which we now turn.

---

[3] It is important to distinguish the *validity* of a principle from its *truth*. Beliefs can be falsely held, but they can also be truly held. Does this mean that the principle is valid in a world where (M)

2.3 Deontic Modalities

Deontic modalities, you will recall from Section 2.1, have to do with permission and obligation. Recall proposition (15): You *can't* torture people.

In a physical sense of 'can', one certainly can torture people. The sense in which (15) is intended to be true is rather a deontic one: it's not *permitted* for you to torture people (or, exploiting the duality from Section 2.1, it is an *obligation* for you not to do so). As with epistemic modalities, and as the plural indicates, deontic modalities constitute a family, unified by the fact that they all relate to permission and obligation. Any code of conduct – in a school, at work, at home, in relation to a game, in a state, in a community, etc. – amounts to a deontic modality. Among the many members in this family, *moral* modality and *legal* modality are most salient. In the moral sense, (15) tells us that torturing people would amount to transgressing a moral principle. In the legal sense, (15) tells us that torturing people is not allowed by a given system of (man-made) laws. Even these types of modalities are, yet again, families: there's not just one system of (man-made) laws in the world, but many; each state (and not only states) has its own set. And, regardless of whether or not there is an absolute (universal) moral code, in an equally legitimate use of the term 'morality', there are various moralities; these are quite often associated with (but not limited to) the variety of religions.

A schema for deontic possibilities can thus be as follows:

p is C-possible if and only if *p* is (logically) compatible with all the elements in C (where C encodes a given code of conduct).

This schema allows us to see a structural difference compared with epistemic modalities: deontic modalities are not relativised to individuals in the way epistemic modalities are.

To give a couple of examples, consider *Christian morality*, according to which you shall not murder, you shall not steal, you shall not commit adultery, etc. We can call 'ChM' the code of conduct encoded in the Ten Commandments. Similar to what we did with epistemic modalities, we can now characterise Christian-Morality modality as follows:

p is Christian-morally possible if and only if it is (logically) compatible with all the elements in ChM.
p is Christian-morally necessary if and only if it (logically) follows from the elements in ChM.

---

doesn't have false instances? The answer is 'no'. It will always be *possible* for it to have false instances, and this possibility *invalidates* the principle qua principle.