What is Direction of Fit?

AVERY ARCHER

ABSTRACT: I argue that the concept of direction of fit is best seen as picking out a certain logical property of a psychological attitude: namely, the fact that it displays a certain two-valued logical structure. Unfortunately, the standard interpretation of the direction of fit metaphor obscures this fact because it conflates two very different properties of an attitude: that in virtue of which it displays a certain direction of fit, and that in virtue of which it displays certain revision-conditions. I claim that the latter corresponds with the aim of an attitude, not its direction of fit. In order to remedy this failure of the standard interpretation, I offer an alternative account of direction of fit, which I refer to as the two-content interpretation.

1. Introduction
What feature or property of an attitude is the notion of direction of fit meant to capture? This is the question I take up in this paper. The answer I arrive at is that the notion of direction of fit is meant to capture a logical property of an attitude. Specifically, it is meant to capture the idea that an attitude displays a certain two-valued logical structure. Unfortunately, the standard interpretation of the direction of fit metaphor obscures this fact because it conflates two very different properties of an attitude: that in virtue of which it displays a certain direction of fit, and that in virtue of which it displays certain revision-conditions. In order to remedy this failure of the standard interpretation, I offer an alternative account of direction of fit, one that succeeds where the standard interpretation fails.

2. The Standard Interpretation
In §32 of Intention, Anscombe employs an example comparing a shopping list and a detective’s record to illustrate two different ways in which our words (written or spoken) may relate to the world. She writes:

Let us consider a man going round a town with a shopping list in hand. Now it is clear that the relation of this list to things he actually buys is one and the same whether his wife gave him the list or it is his own list; and that there is a different relation when a list is made by a detective following him about. If he made the list itself, it was an expression of an intention; if his wife gave it to him, it has the role of an order. What then is the identical relation to what happens, in the order and the intention, which is not shared by the record? It is precisely this: if the list and the things that the man actually buys do not agree, and if this and this alone constitutes a mistake, then the mistake is not in the list
but in the man's performance...; whereas if the detective's record and what the man actually buys do not agree, then the mistake is in the record.\footnote{Anscombe [2000: 56].}

Let us refer to the example discussed here as the \textit{shopping list example}. Anscombe notes that, in the case of the shopping list, the task is to get the world (the items purchased) to match our words (the shopping list), while in the case of the detective's record, the task is to get the word (the detective's record) to match the world (the items purchased). In the terminology of speech act theorists, the former is an example of \textit{world-to-word} fit, and the latter of \textit{word-to-world} fit.\footnote{See and Cf. Searle [1985]).}

Significantly, Anscombe sees an intention (i.e., a certain type of attitude) and an order (i.e., a certain type of speech act) as standing in the same logical relation to the man's actions. By her lights, it does not matter if we see the shopping list as an expression of an intention (on the part of the man) or as an order (on the part of his wife). In either case, the shopping list displays the same direction of fit. Thus, Anscombe takes the observation that our words may relate to the world in two very different ways as also applicable to attitudes. This idea has been taken up by several theorists in moral psychology and the philosophy of mind, where a distinction is often drawn between attitudes that display \textit{mind-to-world} fit and attitudes that display \textit{world-to-mind} fit. Huw Price summarises the central intuition behind the direction of fit analysis of attitudes as follows:

\begin{quote}
Beliefs have a 'mind-to-world' direction of fit; they aim at fitting the world, at being true. Desires, on the other hand have a 'world to mind' direction of fit; we aim to change the world to fit desires, and not vice versa.\footnote{Price [1989: 120].}
\end{quote}

The standard way of unpacking the direction of fit metaphor, which builds on Anscombe's shopping list example, is in terms of whether the relevant attitude or the world is subject to revision when there is a lack of correspondence between the two.\footnote{See and Cf. Platts [1979: 257].} According to this approach, what it means for a belief to have mind-to-world fit is that the belief is subject to revision when there is a mismatch between the world and what is believed. By contrast, what it means for a desire to have world-to-mind fit is that the world is subject to revision when there is a mismatch between that which is desired and the world. I will refer to the preceding interpretation of the direction of fit metaphor as the \textit{standard interpretation}.\footnote{Something along the lines of what I am calling the standard interpretation appears in Platts [1979], Searle [1983], Price [1989], and Humberstone [1992]. For an alternative to the standard interpretation, see Smith [1987].}

Strictly speaking, it is incorrect to say that a belief is subject to revision when there is a mismatch between the belief and the world. After all, if one were to switch from believing $P$ to not believing $P$ (not to be confused with believing $\neg P$), one has essentially given up one's belief, not revised it. What is subject to revision is not the particular token belief but rather the agent's total doxastic state. Whereas the agent's doxastic state initially included the belief that $P$, it is updated so as to no longer include the belief that $P$. However, if we understand the locution “revising a belief” in a broad (albeit somewhat Pickwickian) sense that includes giving it up,
then it seems harmless to say that a belief is subject to revision when it fails to match the world. Moreover, I will continue to speak this way, if only on stylistic grounds. (Talking about the “revision-conditions” of an attitude seems easier on the ear than talk of “giving up-conditions”.) Stylistic considerations aside, the standard interpretation of direction of fit may be summarised as follows: to say that a belief has mind-to-world fit entails that when the world and the belief fail to line up, the burden is on the believing agent to alter her mental representations so as to bring them into line with the world. By contrast, to say that a desire has mind-to-world fit entails that when the world and the desire fail to line up, the burden on the desiring agent is to alter the world so as to bring it into line with her mental representations.

Since many individuals have difficulty keeping tack of the meaning of the locutions “mind-to-world” and “world-to-mind”, I will employ the label theoretical to refer to belief-like mental states, and the label practical to refer to desire-like mental states. The use of the terms theoretical and practical offers the added advantage of having a single label for speech-acts and mental states, and the same direction of fit. For example, we will no longer need to distinguish between word-to-world and mind-to-world fit when referring to the direction of fit of assertions and beliefs respectively. Instead, we can simply apply the label theoretical to both. However, nothing substantive will turn on my choice of terminology. Other labels that have been used in print include cognitive vs. conative and thetic vs. telic. While each of these labels come with their own associations, I aim to cast my discussion at a level of generality that is indifferent to the small variations among these locutions. One should be able to buy into the conception presently on offer irrespective of one’s preferred terminology.

3. Objections to the Standard Interpretation
I believe that the standard interpretation is inadequate for at least two reasons. First, it introduces an unexplained asymmetry between theoretical and practical attitudes. Recall, according to the standard interpretation, saying that an utterance or attitude is theoretical means that it should be withdrawn or given up (respectively) when there is a mismatch between that attitude and the world. However, the standard interpretation is silent about when a practical utterance or attitude should be withdrawn or given up. Instead, we are only left with the negative claim that a practical utterance or attitude is not withdrawn or given up when there is a mismatch between it and the world. Admittedly, it is possible that practical utterances and attitudes simply fail to have revision-conditions. And if this possibility turned out to be an actuality, then it would be unreasonable to require that an account of direction of fit include a specification of the revision-conditions of a practical utterance or attitude. However, such a position seems sufficiently surprising and substantive as to warrant argumentation. It cannot simply be assumed. The result is that the advocate of the standard interpretation faces a dilemma. Either practical attitudes do have revision-conditions analogous to the revision conditions of theoretical attitudes (in which case the standard interpretation owes us an account of the revision-conditions of practical attitudes), or practical attitudes do not have revision-conditions analogous to that of theoretical attitudes (in which case the standard interpretation

---

6 See and Cf.: Humberstone [1992: 60].
owes us an explanation of why practical attitudes lack revision-conditions). Either way, the standard interpretation is, at best, incomplete.

It is worth noting that the present criticism is largely internal to the standard interpretation in the sense that it is an objection that becomes particularly salient given the standard interpretation’s emphasis on the revision-conditions of an attitude. It is not obvious, to me anyway, that we would pre-theoretically expect practical attitudes to have revision-conditions. However, given that the standard interpretation attempts to make sense of the claim that an attitude is theoretical by saying that it has certain revision-conditions, this creates the expectation that an attitude being practical is also to be explained by the fact that it has certain revision-conditions. Hence, the questions of whether or not practical attitudes have revision-conditions, what those revision-conditions consist in (if they do), or why they lack revision-conditions (if they do not), become especially salient given the standard interpretation’s appeal to revision-conditions in its analysis of direction of fit.

Second, the standard interpretation is unable to preserve the intuition that certain attitudes—such as assuming and fantasizing—share the same direction of fit as belief. The notion of direction of fit is most commonly discussed in the context of a contrast between belief and desire.7 However, it is widely held that the distinction is not restricted to belief and desire alone. To say that an attitude has a particular direction of fit is to conceive of it as a member of a broad class of attitudes with which it shares some essential feature. David Velleman puts the point as follows:

> There are many cognitive attitudes other than belief, attitudes that have the same direction of fit and consequently take the same constitutive predicate. Hypothesizing that P, assuming that P, fantasizing that P, and the like are all attitudes in which P is regarded, not as a representation of what is to be brought about, but rather as a representation of what is.8

In the cited passage, Velleman observes that several attitudes, apart from belief, display mind-to-world fit, attitudes that include hypothesizing, assuming, and fantasizing. However, while believing that P and hypothesising that P are plausibly thought of as subject to revision if it is not true that P, the same cannot be said of the attitudes of assuming that P and fantasizing that P. Although assuming that P and fantasizing that P represent things as being a certain way, neither is necessarily subject to revision if P is false. Consequently, if we assume, in keeping with the standard interpretation, that an attitude is theoretical only if it is subject to revision when there is a mismatch between the attitude and the world, then we seem forced to deny that assuming and fantasizing are theoretical attitudes.

The preceding observation suggests a possible diagnosis of where the standard interpretation goes wrong. To say that a particular attitude is theoretical is not yet to say anything about when that attitude is subject to revision. Whether or not an attitude is subject to revision seems tied to its aim, and not to its direction of fit. Hence, a belief is subject to revision

---

7 For a discussion of some of the possible applications of the notion of direction of fit, see Humberstone [1992].
8 Velleman [1992: 12].
if the proposition believed is false because belief aims at truth. Since the attitudes of assuming and fantasizing, by contrast, do not aim at the truth, they are not subject to revision if the proposition assumed or fantasized is false. In light of this, we need to make sense of the direction of fit metaphor in a way that preserves the distinction between saying that an attitude has a certain aim (and by extension, certain revision-conditions) and the claim that an attitude has a particular direction of fit. Unfortunately, the standard interpretation falls short in precisely this respect.

Admittedly, just how much of a challenge the present objection poses to the standard interpretation will depend on how committed we are to saying that assuming and fantasizing are theoretical attitudes. To some, this may be less than obvious. In section 5, I will present considerations that I believe speak strongly in favour of regarding assuming and fantasizing as theoretical attitudes. However, the machinery I will be exploiting in favour of this conclusion has not yet been introduced. In the mean time, suffice it to say that the standard interpretation is unable to preserve the intuition that assuming and fantasizing are theoretical attitudes. Hence, if (like Velleman) one did think that assuming and fantasizing were theoretical attitudes, the present objection would constitute grounds for finding the standard interpretation unsatisfactory. Moreover (and what I take to be the more philosophically interesting criticism), the standard interpretation does not allow for the possibility of two attitudes sharing the same direction of fit if they do not share the same revision-conditions. To the extent that one thinks this is possible (and I present reasons why we should in section 5, below), one should also find the standard interpretation unattractive.

Let us take stock of what we have seen thus far. First, we noted that Anscombe employs the notion of direction of fit to highlight a feature that a desire shares with an order, and that a belief shares with an assertion. This means that a satisfactory account of direction of fit must not only preserve the idea that the notion can be applied to both utterances and attitudes, but that it must offer a univocal account of what direction of fit tracks in both cases. I will refer to this as the univocality requirement. Second, a satisfactory account of direction of fit must preserve the distinction between saying that two (or more) attitudes have the same direction of fit and saying that they have the same revision-conditions. In other words, direction of fit picks out a property that is shared by attitudes that differ in other important ways, including their revision-conditions. I will refer to this as the heterogeneity requirement. Hence, a satisfactory account of direction of fit must satisfy the following two requirements:

**Univocality requirement:**
An account of direction of fit is satisfactory only if it takes the notion to pick out a property that is shared by both a speech act and an attitude.

**Heterogeneity requirement:**
An account of direction of fit is satisfactory only if it allows for the possibility that two attitudes may share the same direction of fit even if they fail to share the same revision-conditions.

---

9 For a discussion of some of the different ways we can understand the claim that belief aims at truth, see Engel [2004].
The standard interpretation satisfies the univocality requirement. Recall, according to the standard interpretation, the assertion that $P$ and the belief that $P$ are both theoretical because they are both subject to revision when $P$ is false. Hence, the standard interpretation offers a univocal analysis of the direction of fit of an assertion (i.e., a speech act) and a belief (i.e., an attitude). However, it fails to satisfy the heterogeneity requirement. The remainder of this paper will be devoted to articulating an alternative to the standard interpretation, one that satisfies both the univocality and heterogeneity requirements.

4. Illocutionary Content

Let us begin our discussion of my proposed alternative to the standard interpretation by defining a few key terms. Typically, an assertion is defined as a speech act in which a proposition is represented as being true. It is widely held that an assertion may share the same propositional content as an attitude, such as believing, assuming or doubting. For example, the assertion that Goldbach's conjecture is true has the same propositional content as the belief that Goldbach's conjecture is true, the assumption that Goldbach's conjecture is true, or the doubt that Goldbach's conjecture is true. Moreover, an agent may use the assertion that Goldbach's conjecture is true to express her belief that Goldbach's conjecture is true. I will refer to an assertion that expresses a belief of the agent making the assertion as a sincere assertion.

There are cases in which an agent makes a putative assertion, but in which the putative assertion does not express one of the speaker's beliefs. I will refer to such putative assertions as insincere assertions. Significantly, to say that a putative assertion is insincere, in the present sense, is not to say that it is somehow infelicitous or inappropriate. Whether an insincere assertion is infelicitous or inappropriate will depend on other factors, such as the context of utterance. For example, if I insincerely assert "I am Harry Potter" in the context of a science fiction and fantasy convention, the fact that my assertion does not express a belief need not render it infelicitous or inappropriate.

An assertion may be codified in a sentence (written, spoken, or signed) by which the asserted proposition is conveyed. I will refer to particular tokens of such written, spoken, or signed sentences as utterances. Two or more different utterances may be used to assert the same proposition. For example, the utterances "It is raining today" and "Es regnet heute" may both be used to assert the proposition [it is raining today]. By necessity, I will be employing written sentences to express the various propositions discussed in this paper. However, it is important to keep in mind that these written sentences will be merely standing proxy for the propositions they are typically used to convey. In order to avoid confusion on this score, I will use double quotation marks to indicate when I'm talking about the utterance "It is raining today" and square brackets to indicate that I am referring to the proposition [it is raining today].

In addition to assertoric utterances, there are also non-assertoric utterances, such as questions, commands, promises and apologies. These all differ in their illocutionary force. However, it is typically assumed that utterances with different illocutionary force may have the same propositional content. For example, consider the following pair of utterances:

---

10 See and Cf. Searle & Vanderveken [1985: 1].
It is widely held that (A) and (B) share the same propositional content; namely, the proposition: [the office door is shut]. The difference between the two utterances has to do with the illocutionary force with which this single proposition is expressed, with the proposition being asserted in (A) and commanded in (B). Since, according to this view, it is possible to vary the force of an utterance while keeping the content fixed, it follows that content and force are independent features of an utterance. I will refer to this view as the independence thesis.\footnote{For a defence of the independence thesis, see Stenius [1967]. See and Cf. Searle & Vanderveken [1985: 1].}

Significantly, the independence thesis relies on a notion of content that leaves the logical properties of an utterance underspecified. This is because a complete characterisation of the logical properties of an utterance—i.e., one that includes the specification of the various deductive inferences for which a particular utterance may be employed—is at least partly determined by the utterance’s illocutionary force. Recall, according to the independence thesis, (A) and (B) both have the same propositional content: the proposition ‘the office door is shut’. If we assume that this content is sufficient to determine the logical character of both utterances, then it would follow that (A) and (B) should be logically interchangeable, despite their contrasting illocutionary force. But this is clearly not the case. For example, an agent who made the following sincere assertions is engaged in a valid piece of inferential reasoning:

\begin{align*}
(A1): & \text{ “The office door is shut.”} \\
(A2): & \text{ “If the office door is shut, then professor Smith is away.”} \\
(A3): & \text{ “Therefore, professor Smith is away.”}
\end{align*}

However, the following putative inference is not valid:

\begin{align*}
(B1): & \text{ “Shut the office door!”} \\
(B2): & \text{ “If the office door is shut, then professor Smith is away.”} \\
(B3): & \text{ “Therefore, professor Smith is away.”}
\end{align*}

The problem with (B1)-(B3) is that the initial premise fails to satisfy the antecedent of the conditional specified in (B2). This leaves us without any basis for inferring the consequent of the conditional, as specified in (B3). Moreover, the failure of (B1) to satisfy the antecedent of (B2) is directly due to its illocutionary force. (B1) fails to satisfy the antecedent of (B2) precisely because it fails to depict the proposition [the office door is shut] as true. In fact, (B1) is perfectly consistent with the falsity of the proposition [the office door is shut]. The same, of course, cannot be said of (A1), which is clearly inconsistent with the falsity of the proposition [the office door is shut]. Thus, although according to the independence thesis, (A1) and (B1) share the same content, (A1) is logically inconsistent with the negation of the proposition [the office door is shut], while (B1) is logically consistent with the negation of the proposition [the office door is shut].
shut]. Given that the only difference between (A1) and (B1) has to do with their illocutionary force, it follows that illocutionary force makes a logical difference.

The takeaway of the preceding discussion is that if we are interested in a complete characterisation of the logical properties of an utterance, we cannot simply consider the utterance’s propositional content. We must also consider its illocutionary force. It will be helpful to have a label for the combination of an utterance’s propositional content and illocutionary force. I will use the label “illocutionary content” for this purpose. In short, the illocutionary content of an utterance is the combination of its illocutionary force and propositional content. Given that the types of deductive inferences in which an utterance can feature depends on the combination of its propositional content and illocutionary force, it eo ipso depends on its illocutionary content.

There are two kinds of illocutionary force that are relevant to the present discussion. The illocutionary force of assertions (e.g., “the office door is shut”) and of commands (e.g., “shut the office door!”). I refer to the former as indicative force and the latter as imperative force. The notion of illocutionary force is paradigmatically applied to speech-acts or utterances. However, I believe the notion may be safely extended to the attitude a speech act or utterance expresses. According to the present suggestion, just as a sincere assertion expresses the attitude of believing, so too a sincere command expresses an attitude of wanting. For example, just as the belief that the door is shut may be expressed by the sincere assertion that the door is shut, so too wanting the door to be shut may be expressed by a sincere command that the door be shut. Hence, I will say that an attitude displays indicative force when it may be expressed by making an assertion, and that an attitude displays imperative force when it may be expressed by issuing a command. Moreover, I will say that an attitude displays indicative illocutionary content when it displays the same illocutionary content as an assertion, and that it displays imperative illocutionary content when it displays the same illocutionary content as a command.

It is important to register that there is something corresponding to what I have been calling the illocutionary content of an utterance or attitude because we assess an agent’s rational standing in light of the illocutionary content of their utterances and attitudes, and not only in terms of the propositional content of their utterances and attitudes. For example, if an agent believes or sincerely asserts that the office door is shut, we take them to have or express an attitude that is inconsistent with disbelieving that the office door is shut. Thus, if an agent were to adopt both an attitude of belief and an attitude of disbelief towards the office door being shut, we would deem them guilty of irrationality. However, consider the case of an agent who wants the office door to be shut. Such an agent has an attitude that may be expressed by sincerely commanding that the office door be shut. Unlike the agent who believes or sincerely asserts that the office door is shut, the agent who wants or sincerely commands that the door be shut has or expresses an attitude that is consistent with disbelieving that the office door is shut. Such an agent is not liable to rational criticism for simultaneously having both attitudes. Moreover, according to the independence thesis, wanting the office door to be shut, sincerely commanding

---

12 Nothing substantive is supposed to hang on my choice of terminology, here. I arrived at the label “illocutionary content” by simply borrowing the word “illocutionary” from illocutionary force and the word “content” from propositional content. This seemed as good a procedure as any for coining a term for the combination of an utterance’s illocutionary force and propositional content.
that the office door be shut, believing that the office door is shut, and sincerely asserting that the office door is shut all have the same propositional content: namely, the proposition [the office door is shut]. Thus, we could not hope to evaluate an agent’s rational standing by simply considering the propositional content of their utterances or attitudes. We must look, instead, to the illocutionary content of an agent’s utterances and attitudes.

5. The Two Content Interpretation

I believe the claim that an utterance or attitude displays a certain direction of fit is a claim about the illocutionary content of that utterance or attitude. Specifically, to describe an utterance or attitude as theoretical is to say that it displays indicative illocutionary content, and to describe an utterance or attitude as practical is to say that it displays imperative illocutionary content. I hold that an utterance or attitude displays indicative illocutionary content just in case its illocutionary content is truth-evaluable. In other words, an attitude displays indicative illocutionary content just in case the combination of its propositional content and illocutionary force yields something truth-evaluable.

Saying that the illocutionary content of believing, assuming, and fantasizing all have truth-evaluable illocutionary content does not entail that their illocutionary content is the same in every respect. One may, for example, think that it is the illocutionary force of a belief that explains the fact that it aims at the truth. If one held such a view, then it would follow that the illocutionary force of a belief must differ from the illocutionary force of assuming or fantasizing, since the latter two do not aim at the truth. Moreover, given that the illocutionary content of an attitude is simply the combination of its propositional content and illocutionary force, then the fact that a belief differ from an assumption and fantasy in terms of its illocutionary force entails that it also differs from them in terms of its illocutionary content. Hence, I do not wish to claim that all three attitudes share the same illocutionary content. Instead, I claim that the illocutionary content of all three attitudes are truth-evaluable. I believe that this is what the claim that believing, assuming, and fantasizing are all theoretical attitudes is meant to capture.

Does it make a difference whether or not we regard assuming and pretending as theoretical attitudes? Given that neither is subject to revision when their content (propositional or otherwise) is false, why should we hold that they have truth-evaluable illocutionary content? It is important to get clear on the nature of the present objection, and on why it poses a putative challenge to my proposed framework. It is relatively uncontroversial that the propositional content of assuming and pretending is truth-evaluable. All propositional content is. (This is equally true of the propositional content of desires, hopes, and wishes.) The salient question is whether what we get when we combine the propositional content and illocutionary force of an assumption or pretending, we are left with something that is itself truth-evaluable. Once we grant that neither is subject to revision when their content is false, and note that this is due to their illocutionary force (which must therefore be importantly different from the illocutionary force of belief), then we seem to be left with little motivation for saying that they have truth-evaluable illocutionary content. In short, why not think that assumptions and pretendings, like desires and wishes, are practical attitudes?

The answer is that, like belief, both assumptions and fantasies can feature in inferences involving truth-evaluable premises and conclusions. This is obvious in the case of assumptions,
since we often use assumptions in the premises of arguments. However, it is also true of pretending. For example, suppose that I am pretending to be Harry Potter (say, while attending a science fiction and fantasy convention), and that my pretence prompts me to make the following set of utterances:

(C1): “I am Harry Potter.”
(C2): “If I am Harry Potter, then Voldemort killed my parents.”
(C3): “Therefore, Voldemort killed my parents.”

By stipulation, (C1) and (C3) do not express beliefs. I do not believe I am Harry Potter; nor do I believe that Voldemort killed my parents. However, (C1)-(C3) still appears to be a perfectly respectable deductive inference for all that. Moreover, (C1)-(C3) conforms to the classical conception of validity; the truth of its premises guarantees the truth of its conclusion. And therein lies the point of contrast between assuming and pretending, on the one hand, and desires, wishes, and hopes, on the other. While the former may be expressed via utterances that conform to a classical conception of validity, the same is not true of the latter.

Recall, according to the two-content interpretation, a desire is expressed via an utterance with imperative illocutionary content. For example, the desire that Parvati shut the door may be expressed via command “Parvati, shut the door!” However, the command “Parvati, shut the door!” is not truth-evaluable. As such, the utterances used to express a desire (or some other practical attitude) do not conform to the classical conception of validity. This follows straightforwardly from the fact that the classical conception of validity only applies to truth-evaluable items. In sum, assuming and pretending are theoretical attitudes because they are expressed with utterances that display truth-evaluable illocutionary content, and which therefore conform to the classical conception of validity.

It is worth emphasising that the preceding argument does not assume that practical attitudes cannot feature in inferences of their own. For example, the following also appears to be an acceptable inference:

(D1): “Fight for freedom or acquiesce to authoritarianism!”
(D2): “Do not acquiesce to authoritarianism!”
(D3): “Fight for freedom!”

(D1)-(D3) appears to be a “valid” inference. However, it cannot be valid in the classical sense, according to which an argument is valid just in case the truth of its premises guarantees the truth of its conclusion. This is because both the premises and conclusion of (D1)-(D3)—being imperatives—are not truth-evaluable. The sense in which (D1)-(D3) is valid seems tied to the fact that obedience to (D1) and (D2) guarantees obedience to (D3). Attempting to capture the sense in which (D1)-(D3) is “valid” is one of the primary goals of an imperative logic.

Given that (D1)-(D3) are imperatives, then it follows from the two-content interpretation that they may express the illocutionary content of a desire, wish, or hope. Consequently, the two-content interpretation allows for the possibility that desires (and other practical attitudes) may feature in “valid” inferences. However, it is also clear that (D1)-(D3) cannot be analysed in terms of a truth-conditional logic. On the contrary, it is the fact that putative “valid” inferences
like (D1)-(D3) exists that gives rise to the need for non-truth-conditional imperative logics. And therein lays the point of contrast between a desire (conceived of as an attitude with imperative illocutionary content) and an assuming or a pretending (conceived of as attitudes with indicative illocutionary content). We do not need to resort to an imperative logic in order to make sense of the fact that an assumption or fantasy may feature in a valid inference. The apparatus of a truth-conditional logic is perfectly adequate for this task. It is the fact that the illocutionary content of beliefs, assumptions, and fantasies may all be captured by a truth-conditional logic that warrants their classification as theoretical attitudes.

The account just adumbrated, while rough, offers us a conception of direction of fit that satisfies both the univocality and heterogeneity requirements. First, it offers a univocal account of what it means for an utterance and an attitude to display a certain direction of fit. Both the speech act of asserting \( P \) and the attitude of believing \( P \) display theoretical direction of fit because they have indicative illocutionary content, and both the speech act of commanding \( P \) and desiring \( P \) display practical direction of fit because they have imperative illocutionary content. As such, the two-content interpretation offers a single explanation of why a speech act or an attitude displays a theoretical or practical direction of fit. Second, it preserves the distinction between saying that an attitude displays a certain direction of fit, and saying that it has certain revision-conditions. According to the two-content interpretation, an attitude is theoretical just in case it has indicative illocutionary content. However, to say that an attitude has indicative illocutionary content is not yet to say anything about when that attitude is subject to revision. Hence, the two-content interpretation allows us to say that believing \( P \), assuming \( P \), and pretending \( P \) are all theoretical attitudes (owing to the fact that they all have indicative illocutionary content) and also say that assuming \( P \) and pretending \( P \) do not aim at truth, and are therefore not subject to revision when they are false. The two-content interpretation therefore satisfies the heterogeneity requirement, at least with respect to theoretical attitudes.

6. Desires, Wishes, and Hopes
In the previous section, we saw that the two-content interpretation satisfies the univocality and heterogeneity requirement in the case of theoretical attitudes. Unfortunately, providing a fully worked out account of practical attitudes would be much too large a task for this paper. However, if an account of practical attitudes is to be satisfactory, then it too must satisfy the univocality and heterogeneity requirements. In the present section, I wish to illustrate that this can be done, within the confines of the two-content interpretation. To this end, I will be sketching an account of practical attitudes. However, my aim is not to argue that the proceeding account of practical attitudes should be accepted. Rather, it is to show that it is possible to offer an account of practical attitudes that both conforms to the two-content interpretation and satisfies the univocality and heterogeneity requirements. This will, I believe, begin to point the way towards the development of a theoretically satisfactory account of practical attitudes, one that takes the univocality and heterogeneity requirements seriously.

In the analysis that follows, I will only be concerned with desire-like attitudes that are directed at an agent or agents, whether the attitude in question is self-directed (e.g., I desire to hit a home run) or other-directed (e.g., I desire that Jackie Robinson hit a home run). This
excludes desire-like attitudes that are not directed at an agent, such as the desire that it be sunny tomorrow or the desire that pi be a rational number. There are three reasons for this restriction. First, the restriction to agent-directed desire-like attitudes is true to Anscombe’s concern with cases of “wanting” that could eventually lead to action. As we noted earlier, Anscombe discusses the concept of direction of fit in the context of speech-acts (or attitudes) that lead to action. Since my wanting it to be sunny tomorrow could not lead to action (either on my part or on the part of others), such instances of wanting seem to fall outside the firing range of Anscombe’s analysis. Second, the restriction to agent-directed desire-like attitudes is in keeping with Anscombe’s suggestion that the notion of direction of fit picks out something that commands share in common with desire-like attitudes. There seems to be a straightforward an intuitive sense in which my desire to shut to door is akin to a self-directed command to shut the door. Likewise, if I desire that someone shut the door, there is an intuitive sense in which such a desire is expressed by a command to shut the door. However, it is less than clear how a desire that is directed at no agent at all is akin to a command. Hence, it is less than clear that Anscombe intended that the notion of a practical attitude be applied to desire-like attitudes that are not agent-directed. Third, it is plausible that agent-directed desire-like attitudes have a different normative significance to those that are not agent-directed. Given that an attitude that is agent-directed is likely to involve norms that fail to apply when no agent is being addressed, it would be presumptuous to assume that a normative treatment of practical attitudes will apply univocally to desire-like attitudes that are agent-directed and those that are not. In light of the preceding considerations, it seems prudent to restrict our analysis of practical attitudes to desire-like attitudes that are directed at an agent.

I hold that like indicative illocutionary content, imperative illocutionary content displays a two-valued logical structure. Arriving at a semantics for the two values in question is one of the primary goals of (non-truth-conditional) imperative logics. My aim, below, is not to contribute to such attempts by offering a formal semantics for imperatives. Rather, it is to provide an intuitive characterisation of the two-valued logical structure of practical attitudes that underscores how they differ in terms of their logical structure from theoretical attitudes. The following intuitive gloss should be adequate for this purpose. Let us say that an utterance or attitude has imperative illocutionary content just in case it demands that some outcome be brought about. Furthermore, let us say that some imperative illocutionary content is doable just in case it is possible for an agent to obey it, and undoable otherwise. On the present view, an utterance or attitude displays imperative illocutionary content just in case its illocutionary content has doability-conditions. In this regard, imperative illocutionary content stands in contrast to indicative illocutionary content, since the former but not latter displays doability-conditions. For example, consider the indicative illocutionary content “the door is shut”. It

---

13 Imperative logics should not be confused with deontic logics since the former require the positing of a different Boolean domain from that employed in classical first-order logic and modal logic, while the latter does not.

14 The notion of possibility I have in mind is not logical or metaphysical possibility, both of which would be much too permissive for our present purposes. Rather, I have something more akin to our quotidian conception of what it means for a course of action to be open to a particular agent. Roughly, is it possible for an agent, S, to bring about an outcome, O, just in case S has the ability and opportunity to bring about O. A great deal more can be said to make the present notion of possibility more precise. However, the preceding intuitive gloss should be adequate for our present purposes.
would be a category mistake to ask if this content can be obeyed, since there is no outcome the agent is being asked to bring about. As such, indicative illocutionary content lacks doability-conditions.

Next, we turn to the differences between the various types of practical attitudes. It is not obvious that there is a sharp distinction between our quotidian conception of a desire and our quotidian conception of a want, hope, or wish. In fact, there are contexts in which the terms are treated as synonymous. For example, saying that one wants to visit Spain, that one wishes to visit Spain, and that one desires to visit Spain may all be taken to express the same attitude. Nevertheless, it does not follow from the fact that a distinction is not already part of our pre-theoretical conception that there is no distinction to be made or worth making. In order to facilitate greater conceptual clarity and ease of philosophical analysis, I propose the following theoretical operationalization of the words “want”, “desire”, “wish” and “hope". I will use the word “want” (and its various cognates) as a generic term for all practical attitudes. I will use the label “desire” to refer to cases of wanting that are directed at outcomes an agent believes she can bring about. In short, a desire is a case of wanting in which the outcome wanted is believed to be doable. According to the present stipulation, one may desire to have a glass of wine (assuming that one believes that having a glass of wine is an outcome one may bring about), but one may not desire to walk across the River Thames (assuming that one does not believe that one can walk across the River Thames).

By contrast, I will be using the label “wish” to refer to a case of wanting that is directed at an outcome that an agent disbelieves she can bring about. According to the present stipulation, an agent may wish to walk across the River Thames (assuming that she disbelieves that walking across the River Thames is an outcome she can bring about), but she may not wish to have a glass of wine (assuming that she believes that having a glass of wine is an outcome she can bring about). Finally, I will use the term “hope” to refer to a case of wanting that is directed at an outcome an agent neither believes nor disbelieves she can bring about. In short, a hope is a practical attitude directed at an outcome an agent is unsure she can or cannot bring about. Hence, an agent may hope to win the race (assuming she neither believes nor disbelieves that winning the race is an outcome she can bring about), but she cannot hope to have a glass of wine, or hope to walk across the River Thames (assuming that these are outcomes she believes and disbelieves she can bring about, respectively).

With the preceding taxonomy of practical attitudes in place, we may now see that the two-content interpretation is also able to satisfy the univocality and heterogeneity requirements in the case of practical attitudes. Desires, wishes, and hopes are all practical attitudes because they display doability-conditions. As such, they all have imperative illocutionary content. Moreover, insofar as the desire that, say, Parvati close the door displays doability-conditions, then so does the command, “Parvati, close the door!”. Hence, the present account offers a univocal account of what it means for an attitude and an utterance to be practical; namely, to possess doability-conditions. The two-content interpretation is therefore able to satisfy the univocality requirement in the case of practical attitudes.

What about the heterogeneity requirement? I believe the preceding account of practical attitudes satisfies this requirement as well. Recall, according to our taxonomy of practical attitudes, what desires, wishes and hopes all have in common is that they all display doability-
conditions. However, saying that a practical attitude displays doability-conditions no more entails that it is doable than saying that a theoretical attitude displays truth-conditions entails that it is true. Just as a theoretical attitude may be false, so too a practical attitude may be undoable. In other words, possessing doability-conditions is as much a prerequisite for an attitude to be undoable as it is for an attitude to be doable. This point parallels the observation that being truth-evaluable is as much a prerequisite for an attitude to be false as it is for an attitude to be true. Moreover, just as claiming that an attitude has truth-conditions does not entail that it is subject to revision if it is not true, so too claiming that an attitude has doability-conditions does not entail that it is subject to revision if it is not doable. For example, if I wish to walk across the River Thames, the object of my wish is not something that is doable. However, since wishing to walk across the River Thames does not involve the belief that doing so is doable, I hold that the wish is not subject to revision because its object is not doable. In this respect, the wish to walk across the Thames is analogous to my pretending that I am Harry Potter. The wish is not subject to revision because what I wish for is not doable and the pretence is not subject to revision because what I pretend is not true.

Given our taxonomy of practical attitudes, only a desire is subject to revision when the outcome wanted is undoable. This is because only a desire aims at the doable. Wishes and hopes, by contrast, don’t aim at the doable, and are therefore not subject to revision when the outcome wanted is undoable. Furthermore, I hold that desires and wishes both have imperative content. As such, it follows from the two-content interpretation that they have the same direction of fit. However, I maintain that a desire is subject to revision if the outcome desired is not something that the agent can bring about directly via the exercise of her own agency. In short, a desire is subject to revision if the outcome desired is not doable. For example, suppose that it is undoable for me to walk across the River Thames. Suppose further that I have the mistaken belief that walking across the Thames is doable. If I desired to walk across the Thames (a desire I could have only if I also had the false belief that walking across the Thames is doable), then my desire would be subject to revision. By contrast, a wish or hope is not subject to revision even if the outcome wished or hoped for is undoable. For example, if I wished that I could walk across the Thames (a case of wanting that counts as a wish because I believe that walking across the Thames is undoable), then the fact that I cannot walk across the Thames does not entail that my wish is subject to revision. A similar argument can be made with respect to the hope that I win the race. Hence, a desire has different revision conditions to a wish or a hope even though all three attitudes have the same direction of fit. Consequently, the two-content interpretation allows us to preserve the distinction between saying that two or more practical attitudes have the same direction of fit and saying that they have the same revision-conditions. It is therefore able satisfy the heterogeneity requirement.

I conclude the present section by noting two limitations of the preceding analysis. First, I do not claim that the notions of the doable and the undoable are analogous to the notions of

---

15 Strictly speaking, I think it gets things wrong to say that a desire has revision-conditions. It would be better to say that desires have correctness-conditions. Nevertheless, I will continue to use the expression “revision-conditions” in order to preserve the terminological continuity between my account of practical attitudes, and my discussion of the standard interpretation. However, those who share my reservations about applying the locution “revision-conditions” to desires may safely substitute the notion of “correctness-conditions” in its place without reducing the force of the arguments below.
truth and falsity in every respect, nor that doability-conditions play the very same role for desire that truth-conditions play for beliefs. (On the contrary, I am of the opinion that nothing plays exactly the same role for desire that truth plays for belief, though I will not attempt to argue for this claim here.\textsuperscript{16}) I exploit the concept of doability-conditions in constructing my account of practical attitudes because it is a way to make sense of the two-content interpretation that satisfies the univocality and heterogeneity requirements. It is therefore important to keep separate the claim that an attitude displays a two-valued logical structure and the claim that the two values in question are analogous, in every respect, to truth-values. Second, I am not claiming that the preceding analysis of practical attitudes is one we should adopt, at least in its present form. What I have attempted to do is demonstrate that it is possible to arrive at an account of practical attitudes that satisfies the requirements set forth in this paper, and illustrate what it might look like. Moreover, I believe that the preceding analysis may at least serve as a starting place for developing a more nuanced and sophisticated account of practical attitudes, one that (i) offers a univocal account of what makes commands (i.e., a certain kind of speech-act) and desires (i.e., a certain kind of attitude) practical, and (ii) respects the difference between saying that a practical attitude has certain aim (if any), and saying that it displays a certain direction of fit.

7. Conclusion
In this paper I have attempted to get clear on what property of an attitude the notion of direction of fit is meant to pick out. According to the standard interpretation, the notion of direction of fit picks out the property in virtue of which an attitude has certain revision-conditions. I claim that this is a mistake. An attitude has the revision-conditions that it does in virtue of its aim. It is the fact that belief aims at the truth that accounts for the fact that it is subject to revision if it is false.\textsuperscript{17} The mistake of the standard interpretation lies in its failure to register that there is a gap between saying that an attitude is truth-evaluable, and saying that it is subject to revision if it is false. Attitudes like assuming and pretending inhabit this gap. Contra the standard interpretation, I claim that the notion of direction of fit picks out the property in virtue of which an attitude displays a certain two-valued logical structure. When we conceive of the notion of direction of fit along these lines, we are able to make room for not only believing and hypothesizing, but also for assuming and pretending, as theoretical attitudes. According to my proposed account, what all theoretical attitudes have in common is that they may feature in valid deductive inferences, in the classical sense of validity. By contrast, practical attitudes cannot feature in classically valid deductive inferences. It is this point of contrast between theoretical and practical attitudes that I believe the notion of direction of fit is meant to capture.

\textsuperscript{16} For arguments along these lines, see Velleman [1992].
\textsuperscript{17} Of course, one may still choose to use the term “direction of fit” to refer to the very property that the notion of the aim of an attitude already tracks. However, this would be to make the notion of direction of fit redundant. By contrast, my conception of direction of fit preserves its status as a theoretically interesting notion in its own right.
References


Stenius, E. [1967], “Mood and Language-Game”, *Synthese*, 17: 254-274.