Reasoning in attitudes

December 2022

final version appearing in *Synthese*[1]

Franz Dietrich
Paris School of Economics & CNRS

Antonios Staras
University of Cardiff

Abstract

People reason not only in beliefs, but also in intentions, preferences, and other attitudes. They form preferences from existing preferences, or intentions from existing beliefs and intentions, and so on. This often involves choosing between rival conclusions. Building on Broome (2013) and Dietrich et al. (2019), we present a philosophical and formal analysis of reasoning in attitudes, with or without facing choices in reasoning. We give different accounts of choosing, in terms of a conscious activity or a partly subconscious process. Reasoning in attitudes differs fundamentally from reasoning *about* attitudes, a form of theoretical reasoning in which one discovers rather than forms attitudes. We show that reasoning in attitudes has standard formal properties (such as monotonicity), but is indeterministic, reflecting choice in reasoning. Like theoretical reasoning, it need not follow logical entailment, but for a more radical reason, namely indeterminism. This makes reasoning in attitudes harder to model logically than theoretical reasoning. But it can be studied abstractly, using indeterministic consequence operators.

1 Introduction

A growing philosophical literature about rationality and reasoning teaches us that beliefs are not the only locus of reasoning. You also reason in intentions, preferences, and other attitudes (e.g., Broome 2006, 2013, Kolodny 2005, 2007, Boghossian 2014, Dietrich et al. 2019). By reasoning you can form preferences from preferences; this makes your preferences more transitive. You can form the intention to help a child cross a street from believing you ought to; this makes you less akratic. You can form the same intention from intending to make the child happy and believing that reaching this end requires your help; this makes you more instrumentally rational. And so on.

Two questions guide this paper:

[1] We are grateful for rich feedback and suggestions from John Broome, Christian List, Robert Sugden, Frederik van de Putte, other colleagues, and referees. The paper subsumes our earlier working paper with the main title "Reasoning in versus about attitudes". Franz Dietrich acknowledges support from the French Research Agency through the grants ANR-17-CE26-0003, ANR-16-FRAL-0010 and ANR-17-EURE-0001.

(1) What *is* reasoning in attitudes, philosophically and formally?
(2) Can such reasoning follow entailment of some type, and hence be modelled logically?

On question (1): For one, reasoning in attitudes – which John Broome calls reasoning 'with' attitudes[2] – differs fundamentally from reasoning about attitudes, in which you discover rather than form attitudes. Reasoning in attitudes is an internal process of creating attitudes by first-personally entertaining attitudes, i.e., activating existing attitudes and forming new ones. It happens routinely in daily life. An important example is practical reasoning, in which you create intentions underlying actions. It matters directly to practical philosophy, psychology, and even artificial intelligence.[3] By contrast, reasoning about attitudes is an internal process of discovering attitudes by third-personally observing attitudes, i.e., recalling the existence of some attitudes and deducing the existence of others. Here the attitudes in question could be your own attitudes (in an act of reflection) or someone else's attitudes (e.g., another player's intentions in a game; cf. Perea 2012). Such reasoning is a form of theoretical reasoning, i.e., reasoning in beliefs. Specifically, it is theoretical reasoning *about attitudes*.

For another, reasoning in attitudes differs from other processes that also change your attitudes, including processes driven by external causes (music can create desires) and internal psychological processes that are purely automatic and unconscious (desires can cancel intentions that stand in the way). We focus exclusively on *reasoning*.

Our philosophical account of reasoning in attitudes starts from Broome's (2013) influential work, but presents a novel analysis of choice in reasoning that shows how choosing is possible and is inextricably related to practical reasoning. Our Broome-inspired take on reasoning contrasts with more liberal approaches that count far more mental processes as reasoning.[4] Our formalisation of reasoning draws on Dietrich et al. (2019) and on the apparatus of abstract consequence operators.

On question (2): As we shall see, reasoning in attitudes goes fundamentally beyond entailment, even when twisting the notion of entailment. The main reason is choice in reasoning: when 'choosing' a conclusion, you reason to something that is optional and hence not entailed.

The difficulty to model reasoning in attitudes logically might surprise at first. After all, logic is very useful when modelling (deductive) *theoretical* reasoning;[5] and logic provides powerful tools to model attitudes, namely modal operators such as belief operators, preference operators, or intention operators.[6] One might thus have conjectured that logic can 'somehow' model reasoning in attitudes, modulo standard idealisations or abstractions that come with any formal model. Truth of this conjecture is implicitly

[2] Our terminology aims to prevent any confusion with 'reasoning about attitudes'.

[3] Sophisticated intelligent systems use (artificial) reasoning to form (artificial) attitudes, including intentions that cause actions.

[4] An example is Drucker's (2021) broader account of reasoning, which he calls 'generalism'.

[5] Non-deductive reasoning can behave very differently, for instance non-monotonically. See Harman's (1984) seminal distinction between reasoning (inference) and entailment (implication).

[6] Examples of logics with attitude operators are logics of preferences (e.g., Liu 2011), of beliefs (e.g., Halpern 2017), or of beliefs, desires and intentions ('BDI logics').

presupposed by the common conception of an ideal reasoner as someone whose set of attitudes is closed under entailment. If, as we argue, an (ideal) reasoner must sometimes 'choose' which of certain attitudes to derive, none of which is entailed, then closure under entailment is not the characteristic mark of (ideal) reasoning. Having a closed set of attitudes might be necessary for being an (ideal) reasoner, but not sufficient.

Our analysis will reject the conjecture. This does not mean that reasoning in attitudes cannot be modelled: such reasoning can (and will) be modelled using the abstract machinery of consequence operators, simply because consequence operators can represent *any* transformation of attitudes. Our doubts about modelling reasoning in attitudes refer to the use of concrete logics, not abstract logic.

2 Reasoning in attitudes – deterministic case

This section discusses and formalises reasoning in attitudes without yet addressing choice in reasoning (which we treat in Section 3). The philosophical account follows Broome (2013), and the formalism follows Dietrich et al. (2019).

2.1 Attitudes and constitutions

The agent – 'you' – holds various attitudes, such as: believing that it snows, desiring to feel warm, intending to dress warm, preferring snow to rain, etc. Formally, we fix a set M of possible *attitudes* or *(mental) states*. Those attitudes that you possess form your *constitution*:

Definition 1 *A **(mental) constitution** is a set $C \subseteq M$ of attitudes ('your' attitudes).*

Think of attitudes in M as pairs of an attitude-content and an attitude-type. For many philosophers, contents are propositional: they are *single* propositions for monadic attitudes like intention, *pairs* of propositions for dyadic attitudes like preference, etc.[7] One could make this structure of states formally explicit.[8]

We use the term 'attitude' not only for mental states in M (such as: desiring to be warm), but also for attitude-types (such as: desire).

2.2 The Broomean account in a nutshell

Your constitution changes through reasoning. In reasoning, you form a (conclusion-)attitude from existing (premise-)attitudes: you form beliefs from beliefs, intentions from beliefs and desires, preferences from preferences, etc. The process is causal: the premise-attitudes cause the conclusion-attitude. It constitutes a conscious mental act.

[7] On 'propositionalism', see Felappi (forth.). For a critical take, see Montague (2007).

[8] Let L be a set of *propositions*, and A a set of *attitude-types*, each carrying an *arity* $n \in \{1, 2, ...\}$, usually 1 (monadic attitudes) or 2 (dyadic attitudes). Plausibly, A contains at least belief *bel* (monadic), desire *des* (monadic), intention *int* (monadic), preference $\succ$ (dyadic), and indifference $\sim$ (dyadic). Finally, define *attitudes* in M as tuples $m = (p_1, ..., p_n, a)$ where a is an attitude type in A, n is its arity, and $p_1, ..., p_n$ are propositions in L. So, (p, bel) is believing p, (p, int) is intending p, $(p, q, \succ)$ is preferring p to q, etc.

This complex act consists in bringing the premise-attitudes to mind and then forming the conclusion-attitude. If you reason explicitly (as we shall mostly assume), you say the contents of all attitudes to yourself, often using silent internal speech. You might reason:

Paying taxes is legally required. So, I shall pay taxes. (1)

This is reasoning from a single premise-attitude, namely a belief, to an intention. The conclusion-attitude has the content *I will pay taxes.* What you say however involves 'shall', here treated as a linguistic marker indicating that you entertain the content as an intention.[9] In reasoning, you express to yourself the *marked contents* of your premise- and conclusion-attitudes, not the contents simpliciter. Marked contents are contents marked by 'how' the content is entertained: as a belief, or intention, etc. The English language contains markers for various attitude types, allowing you to reason in those attitudes. Beliefs are special in that they have no linguistic marker, in other words have a 'silent marker'. The same sentence – in the example: *Paying taxes is legally required* – expresses the belief's content and its marked content.

It is debatable how exactly the English language expresses reasoning, i.e., which linguistic constructs serve to mark attitude-contents. Reasoning in preferences might at first seem obscure, as preferences are dyadic attitudes. Broome (2006) however points out (citing Jonathan Dancy for this insight) that English has a preference marker, namely a construction with 'rather'. You can reason in preferences as follows:

Rather bike than walk. Rather walk than drive. So, rather bike than drive. (2)

You initially prefer biking to walking, and waking to driving. You come to prefer biking to driving.

Crucially, you do not say to yourself *that you hold* the attitudes in question. For instance, in (1) you do not say:

I believe paying taxes is legally required. So, I intend to pay taxes.

This would be reasoning *about* your attitudes (cf. Section 4.2).

Reasoning is rule-governed: you draw the conclusion by following a *rule*. A rule is something that allows forming a (conclusion-)attitude from existing (premise-)attitudes. What is it exactly? For now we focus on deterministic rules, i.e., rules in Broome's sense.

A rule can be more or less specific. The most specific rule you can follow in (1) is this: from believing that paying taxes is legally required, come to intend to pay taxes. You could follow a broader rule, which is representable as a schema, such as: from believing that ϕ-ing is legally required, come to intend to ϕ (where ϕ is any act).

Whether real agents reason with specific or broad rules is a deep question that we cannot settle. But we add three speculative thoughts. First, reasoning with broad rules

[9] Using 'shall' to mark intention is little common in everyday English, although some American philosophers have appropriated 'shall' as a marker for intention, as a referee pointed out to us. In original English, almost the opposite used to hold: 'shall' was the correct auxiliary of the future tense after 'I' or 'we', and saying 'will' instead of 'shall' counted as a departure that adds extra colour and might mark an intention. In modern English, 'will' has mostly taken the role of 'shall'.

and reasoning with specific rules is *effectively* equivalent, in the following precise sense: to any broad rule correspond certain specific rules (its instances), and the conclusion-attitudes derivable using certain given broad rules are precisely the conclusion-attitudes derivable using the corresponding specific rules.[10] So, the fact of whether you reason with broad or specific rules is underdetermined by the effects of reasoning on attitudes, i.e., by the fact of which new attitudes you (can) create by reasoning. Second, this sort of underdetermination does not make the nature of your rules indeterminate: there may be a psychological fact about the nature of your rules, whether or not this fact has attitudinal implications. Third, the idea of non-specific rules raise a problem: such rules involve a parameter or placeholder (the 'ϕ' in our example), and it is not clear how the conscious mental act of reasoning can handle such such 'anonymous' parameters and their precise domains. This difficulty would perhaps be insurmountable if you (the reasoner) had to be aware of the rule you follow, or even had to spell it out. But, fortunately, a reasoner need not be aware of the rule, or even of the concept of rule. Even when reasoning explicitly, you only spell out your premises and conclusion, not your rule.

Many rules promote your rationality. Here are examples of rationality-promoting rules, stated informally:

(a) *Modus-Ponens Rule*: From believing p and believing *if* p *then* q, come to believe q. Parameters: propositions p, q.

(b) *Enkratic Rule*: From believing *obligatorily* p, come to intend p. Parameter: propositions p.

(c) *Instrumental-Rationality Rule*: From intending p and believing q *is a means implied by* p, come to intend q. Parameters: propositions p, q.

(d) *Preference-Transitivity Rule*: from preferring p to q and preferring q to r, come to prefer p to r. Parameters: propositions p, q, r.

For instance, your reasoning in preferences (2) follows the Preference-Transitivity Rule, where p, q and r are *I bike*, *I walk* and *I drive*, respectively. The rules (a)–(d) are just examples; one could refine these rules (e.g., by reformulating or adding premises) or state entirely other rules. What the right rules are is not our topic.

2.3 Reasoning formalised – deterministic case

We now formalise Broomean reasoning following Dietrich et al. (2019), still ignoring choice in reasoning.

For simplicity, our model uses rules of specific rather than broad type. Technically speaking, this is not a true restriction of generality, because readers who think that the real reasoner follows broad rules can reinterpret the model's rules (of specific type) as instances of the real rules (of broad type).

[10] For instance, assume you have just one rule: the broad rule that takes you from believing that the weather is X to intending to take a walk, where X is 'dry' or 'windy'. Having this rule is effectively equivalent to having two specific rules: the rule taking you from the dry-weather belief to the intention and the rule taking you from the windy-weather belief to the intention. Indeed, whether you have the broad rule or the two specific rules, you can derive the intention if and only if you initially have one of the two beliefs.

So, we can define a (deterministic) reasoning rule simply as a combination (P, k) of a specific set of (premise-)attitudes $P \subseteq M$ and a specific (conclusion-)attitude $k \in M$. The four rule schemas (a)–(d) in Section 2.2 can now be re-stated more formally:

- the rules $(P, k) = (\{\textit{believing}\, p, \textit{believing}\, \textit{if}\, p\, \textit{then}\, q\}, \textit{believing}\, q)$ for propositions p, q,
- the rules $(P, k) = (\{\textit{believing obligatorily}\, p\}, \textit{intending}\, p)$ for propositions p,
- etc. for (c) and (d).

These re-statements of rules stay semi-informal; but formal statements are possible.[11]

A rule (P, k) transforms your initial constitution C. If you possess all premise-attitudes in P but not the conclusion-attitude k, then you form the attitude k, acquiring the constitution $C \cup \{k\}$. Otherwise your constitution remains the same. Formally:

Definition 2 *A **deterministic reasoning rule** is a pair (P, k) of a set of premise-attitudes $P \subseteq M$ and a conclusion-attitude $k \in M$. The **revision** of a constitution C by this rule is the constitution*

$$C^+ = \begin{cases} C \cup \{k\} & \textit{if } P \subseteq C \ \& \ k \notin C \ \textit{(the rule 'applies')} \\ C & \textit{otherwise (the rule does not 'apply').} \end{cases}$$

You can reason with certain rules – 'your' rules. The set of your rules is your *reasoning system* S. A constitution is *reachable* for you if you can acquire it by reasoning with any number of your rules. Formally:

Definition 3 *A **deterministic reasoning system** is a set S of deterministic reasoning rules ('your' rules). Given an initial constitution C, a constitution C^+ is **reachable from C by (S-)reasoning** if there are a number of steps $n \geq 0$ and constitutions $C_0, .., C_n$ such that $C_0 = C$, $C_n = C^+$, and each C_t with $t \neq 0$ is the revision of C_{t-1} by some rule in S.*

3 Reasoning in attitudes – indeterministic case

You often face choices in reasoning. Broome (2013) discusses such choice, but develops no full account. We now present a philosophical analysis of indeterministic reasoning in attitudes (Sections 3.1–3.4) and a formal model thereof (Sections 3.5–3.6). Philosophically oriented readers can focus mainly on Sections 3.1–3.4, technically oriented readers mainly on Sections 3.5–3.6. An appendix adds complementary discussion and formal material.

[11] Using the formalism in footnote 8, write $(P, k) = (\{(p, bel), (\textit{if}\ p\ \textit{then}\ q, bel)\}, (q, bel))$ $(p, q \in L)$ for (a); write $(P, k) = (\{(\textit{obligatorily}\ p, bel)\}, (p, int))$ $(p \in L)$ for (b); etc. This involves composite propositions. To give them formal meaning, assume that to any propositions $p, q \in L$ is assigned a proposition $\textit{if}\ p\ \textit{then}\ q$ in L; that to any proposition $p \in L$ is assigned a proposition *obligatorily* p in L; etc. Technically, this defines a binary operator $L \times L \to L$; a unary operator $L \to L$; etc. The rules (a)–(d) are now formally specified. One could go further and model propositions in L syntactically (intensionally) as sentences in a formal language, or semantically (extensionally) as subsets of some set of possible worlds. This turns operators into syntactic or semantic operators, respectively (cf. Dietrich et al. forth.).

3.1 The problem of choice in reasoning

Consider a version of an example in Broome (2013: 163 ff.). You intend to visit Venice; you believe that taking a boat and taking a train are the two only possible means; and you believe that both means are equally good. From these three premise-attitudes you can reason either to a 'boat' intention or to a 'train' intention. You face choice in reasoning. Facing choice happens frequently in instrumental reasoning – not just when you believe that the available means are equally good, but also when you are unable to compare the means, and more generally when your attitudes create a 'tie'.[12]

We see two fundamental types of choice in reasoning, to be called 'choice-between' and 'choice-whether'. The former is a choice between rival conclusions, as in our example. The latter is a choice as to whether to derive any conclusion at all. You can face a choice-whether without a choice-between (here there is just one conclusion, which is optional), or a choice-between without a choice-whether (here there are many possible conclusions, abstention being impermissible), or both simultaneously. Broome's account of reasoning (Section 2.2) already allows for choice-whether, in that it does not imply a *requirement* to reason with a given rule – or so Broome would argue. We shall thus focus on choice-between (but return to choice-whether in the conclusion). 'Choice' and 'indeterminism' will therefore usually refer to choice-between.

The question of how you reason when facing a choice between rival conclusion-attitudes contains two subquestions:

(i) How do you reason to any conclusion-attitude at all?

(ii) How do you avoid reasoning to several conclusion-attitudes consecutively?

Both questions are non-trivial. In (i), the possibility to reach a conclusion *by reasoning* is prima facie threatened by the lack of reasons for favouring any conclusion over another. In (ii), the difficulty is that, prima facie, *if* you can reach some conclusion *then* you can by symmetry reach another conclusion too, assuming that your premise-attitudes are all you need to reach a conclusion.

Broome's (2013) take on question (i) is strikingly simple: you can reason to a conclusion in the same way in which you do it when you face no choice, i.e., when there is no alternative conclusion. Formally speaking, if in the Venice example your initial constitution C includes the set P of premise-attitudes, and if your reasoning system contains the deterministic reasoning rules (P, k_1) and (P, k_2), where k_1 and k_2 are the possible conclusion-attitudes, i.e., the 'boat' and 'train' intention, then you can form either conclusion by standard deterministic reasoning. We shall adopt this approach to question (i), but indicate an alternative in the concluding remarks.

We shall grapple much with question (ii). The trouble is that nothing in Broome's account stops you from forming 'surplus conclusions', as Dietrich et al. (2019) note. You can form both intentions by applying both rules (P, k_1) and (P, k_2) consecutively; and this *will* happen if you reason all the way, applying all your rules. What mechanism stops this counterintuitive 'surplus reasoning'?

[12] In the Venice example we could have replaced your equal-goodness belief (the third premise-attitude) by another tie-creating attitude, e.g., a preferential indifference between the means, or a belief of being unable to compare the means in terms of betterness or equal goodness.

3.2 Four accounts of indeterministic reasoning

We now present four accounts of how you make *one and only one* choice in reasoning. Technically, a situation of choice in reasoning is characterised by a set of premise-attitudes $P \subseteq M$ (in the Venice example: the intention to visit Venice and two beliefs) and a set of at least two possible conclusion-attitudes $K \subseteq M$ (in the Venice example: the 'boat' intention and the 'train' intention). Each account of choice in reasoning gives *some* explanation of how you succeed in reasoning from P to one and only one attitude in K.

We use the Venice example to illustrate each account. So, we must explain how you reach the 'boat' or 'train' intention, but not both. In principle, the four accounts generalise to other examples, in which the conclusion-attitudes can be of other types than intentions and of another number than two. Still the Venice example is paradigmatic, as it addresses action-guiding (practical) reasoning and as it parallels the infamous problem of Buridan's ass.

First account. This account decomposes the process into two stages. First you reason to the broad intention to take either a boat or a train. At some later point, a psychological process (discussed below) refines your intention, creating either the 'boat' intention or the 'train' intention. Your inner speech might be as follows, with comments added in square brackets:

> *I shall visit Venice. For this I must either take a boat or take a train.*
> *Both are equally good. So, I shall take a boat or a train.* (3)
> [A psychological process intervenes.] *I shall take a boat.*

The final sentence, which expresses your refined intention, does not start with 'So', as we assume here that it is not formed by reasoning. In a variant of (3), the refinement of your intention is not verbalised, i.e., your inner speech act stops after forming the broad intention, which is particularly plausible if the refinement process is subconscious.

What kind of process refines the intention? We see two possibilities.

(A) *Automatic refinement.* Here an automatic (and possibly subconscious) process refines your intention. This happens routinely in life, especially in trivial situations such as when your intention to buy a bottle of wine gets refined automatically (as you enter the shop) into the intention to buy the bottle on the left, caused by being left-hand or by seeing this bottle first. Surplus intentions are avoided by the (plausible) assumption that your automatic psychology creates only one specific intention.

(B) *Active refinement.* Here you 'do' the refinement actively. You actively (and consciously) form the 'boat' intention or 'train' intention. What sort of act might this be?

  (B1) *Either* your act of refining the intention is a (second) act of reasoning: you derive a specific intention, premised on your broad intention and your belief that both means are equally good. Your speech then differs slightly from (3), as you start the final sentence with 'So'.[13] This makes the first account of

[13] You presumably do not repeat the premises of your second reasoning because you have just said

indeterministic reasoning circular: indeterministic reasoning is decomposed into two reasoning steps, the second one being again indeterministic. The choice has stayed exactly the same: a choice between two intentions. Only the premises have changed. Both fundamental problems of indeterministic reasoning – How do you choose a conclusion? How do you avoid choosing a second one afterwards? – reemerge at the second stage. The account can be made non-circular by adding an explanation of the indeterministic reasoning at stage two. This explanation might take the form of one of our three later accounts, suitably adapted to the new premises. It is not clear prima facie why choosing a conclusion should be suddenly possible at stage two if it was not possible earlier based on the initial set of premises P.

(B2) *Or* your act of refining your intention is a 'mere choice', distinct from reasoning. You choose a specific intention, without deriving it from premises. You simply decide to take a boat, or (in our wine example) to buy the bottle on the left. This idea is strikingly simple. Does it work? You presumably make this choice on a basis. The basis is, it seems, that you initially intend to take a boat or train (which justifies forming a specific intention) and that you believe that both means are equally good (which justifies forming *any* of the specific intentions rather than having to form the 'boat' intention or having to form the 'train' intention). These two attitudes, on which your 'mere choice' is based, were precisely the premise-attitudes when refining by reasoning in (B1). This lets your 'mere choice' in (B2) look dangerously similar to your reasoning in (B1). The fact that a mere choice is likely to be done implicitly need not make a compelling difference to reasoning, since reasoning can also be implicit. Does the hypothesis of refinement by mere choice in (B2) therefore collapse into that of refinement by reasoning in (B1)? If not, what distinguishes the two? We leave these questions open.

On a generalised version of the first account, the broad conclusion need not be refined in one go, but can be refined gradually: you first move from the broad conclusion to a finer one, then to an even finer one, and so on, until reaching a specific conclusion in K. Each step can be automatic or active. Each step eliminates some member(s) of K, until just one member is left. In our Venice example the refinement must go in one step, as K has only two members, hence only one member to eliminate.

On the first account of indeterministic reasoning from the premise set P to an attitude in K, such reasoning is only apparent: you do not actually reason from P into K, but you undergo a two-step process, in which only the second step is indeterministic, and that step neither starts from P, nor is necessarily a *reasoning* step. For these reasons, one might reject the first account on the 'technical' grounds of being an account of something else than reasoning from P into K. However, on more substantive grounds, the process envisioned by the account is clearly realistic. You often begin by deriving a broad conclusion, which is refined later. But, equally clearly, you often reason directly from P to a specific conclusion in K, without taking the detour over a broad conclusion.

them during your first reasoning.

This is indeterministic reasoning from P into K in a *literal* sense, unlike in the first account. The next three accounts explain how this could work.

Second account. Here we presuppose a wider notion of reasoning: you can reason not just from attitudes, but also from absences of attitudes. You reason to the 'boat' intention or the 'train' intention based on your three attitudes *and* on the two additional premises of not yet having this intention and not yet having the other intention (or at least the one additional premise of not yet having the other intention). This blocks surplus reasoning because after reasoning to one intention – say the 'boat' intention – this intention is no longer absent, so that one additional premise no longer holds. No broad intention is needed under this account.

Yet reasoning from absences violates the spirit of our Broomean account of reasoning, to say the least. First, it is not explicit. You can say to yourself *that* some attitudes are absent, but this expresses a belief in the absence, not the absence itself. It reports rather than expresses the absence. Your inner speech on this account might be this (note that both 'premises of absence' remain unexpressed):

> *I shall visit Venice. For this I must either take a boat or take a train. Both are equally good.* [The absence of the 'boat' intention and the 'train' intention intervene implicitly.] *So, I shall take a boat.*

Second, reasoning from absences is not, or not in the same way, conscious. Why? An ordinary Broomean reasoner brings to mind the premise-attitudes, thereby making them conscious. But you cannot bring to mind an absence (what would that mean?). If anything, you can bring to mind the belief in that absence. And bringing this meta-belief to mind might not even be possible, as you might not *have* it. Indeed, reasoning from absences requires only that you do not possess certain attitudes, not that you believe (know) that you do not possess them – just as standard Broomean reasoning requires only that you possess the premise-attitudes, not that you believe (know) that you possess them.

Third, against this background reasoning from absences is probably not an act: it is not something you do.[14]

Third account. This account might be the one that is most in the spirit of Broome's (2013) short discussion of choice in reasoning. On the account, after reasoning to either intention, say the 'boat' intention, you *can but do not* reason to the 'train' intention. You *can* reason to the boat intention because your three premise-attitudes remain present and are all you need to derive a 'train' intention, unlike under the second account. But you *do not* derive the 'train' intention. Broomean reasoning is something you can, but need not do. Your premise-attitudes enable, but do not require forming the conclusion-attitude.

This begs a question: how do you avoid reasoning again? Here the account goes beyond Broome (2013), filling a gap. On the account, your 'boat' intention prevents the

[14] If reasoning from absences is not explicit, and only explicit reasoning can be active (as suggested but not assumed by Broome 2013, p. 224), then reasoning in absences is not active.

second reasoning from getting started or completed. From the three premise-attitudes *plus* your 'boat' intention you can no longer derive the 'train' intention, because deriving the 'train' intention from the (now) *four* premises seems wrong to you, hence does not constitute Broomean reasoning.

One might object that, despite having the 'boat' intention, you might not bring it to mind. You might only bring to mind the three original premise-attitudes, from where you *can* derive the 'train' intention. Yet, on the account, this simply never happens: you always bring your 'boat' intention to mind, if it is not already conscious. More precisely:

(i) If your 'boat' intention is already conscious, perhaps because you have just acted on it by booking a boat ticket, then you usually do not even start reasoning to another intention. And even if you did start, i.e., brought to mind the three premise-attitudes, you could not complete your reasoning. You could not derive the 'train' intention from the *four* attitudes you now have in mind, including the 'boat' intention.

(ii) If your 'boat' intention is initially subconscious ('inactive'), perhaps because you formed it too long ago, then you might begin to reason. You bring to mind premises. But then, rather than rushing to the conclusion of a 'train' intention (by Broomean reasoning), you remember your 'boat' intention, by bringing it to mind ('activating' it). You remember it because your activity of reasoning makes it salient. You remember it after bringing all three premise-attitudes to mind, or already after bringing some of them to mind. Then you stop reasoning, again because you cannot derive a 'train' intention from the attitudes you now have in mind, including your 'boat' intention.

Your inner speech act might be as follows, assuming case (ii) and assuming you remember your 'boat' intention after bringing to mind two premises:

> *I shall visit Venice. For this I must either take a boat or take a train.*
> *Both are equally good. So, I shall take a boat.*
> [Break. Intentions and beliefs become unconscious. You start again:] (4)
> *I shall visit Venice. For this I must either take a boat or take a train.*
> *Oh, but I already intend to take a boat.* [Reasoning stops.]

In another version of the third account, specifically of clause (ii), you remember your 'boat' intention because it *comes* to mind *automatically*, not because you *bring* it to mind *actively*. The effect is the same: you do no longer derive the 'train' intention.

Fourth account. Perhaps the third account is satisfactory. But arguably a problem remains: reasoning from the three premises to intending a means is arguably *incorrect* reasoning, against what Broome thinks.[15] To derive the intention *correctly*, you need an extra premise-attitude, for instance the belief that you do not (yet) intend either means,

[15] More precisely, Broome regards such reasoning as correct provided one adds a fourth premise-attitude, namely the belief that both means are up to you. If the means are not up to you, say because you have delegated your travel planning to a travel agent, then your intention is unnecessary. Broome may be right that correctness requires his fourth premise. We ignore this premise for expositional simplicity.

or the belief that you do not (yet) intend the other means, or perhaps something else. These claims about correctness assume Broome's own notion of correct reasoning. They are discussed in Appendix B; here we simply accept them.

The fourth account adds the missing fourth premise-attitude. For concreteness, let this attitude be the belief that you do not yet intend either means (some other versions would also work, mutatis mutandis). This additional belief is introspective or 'second-order' or 'meta': it is a belief about your attitudes. This premise-belief often remains implicit, and possibly subconscious. But even when it is implicit or subconscious, it needs to be present, and *could* be made explicit. Sometimes this belief is initially absent, and is formed while reasoning to an intention. It may be formed actively, as discussed in Section 3.4.

The speech act by which you form the 'boat' intention might be this:

> *I shall visit Venice. For this I must either take a boat or take a train. Both means are equally good.* [Your meta-belief is formed, actively or automatically.] *I do not yet intend a means. So, I shall take a boat.*

Appendix A presents a more complete version of your inner speech in case your meta-belief is formed actively.

Curiously, your reasoning makes one of its premises false. Indeed, the premise-belief of not intending either means becomes false by forming either intention, say the 'boat' intention. After becoming false, the belief might disappear automatically. This prevents you from reasoning to the 'train' intention, as a premise has disappeared. But, even if your introspective belief fails to disappear after becoming false, then something else saves you from starting or from finishing to reason to a 'train' intention: this reasoning is blocked by your 'boat' intention, which is either already conscious, or is brought to mind actively, or appears automatically. The precise mechanism parallels what we encountered in the third account.[16]

3.3 Indeterministic reasoning as an activity

Indeterministic reasoning is not standard Broomean reasoning, because it is an interplay between some process generating a conclusion-attitude and some mechanism blocking a second conclusion-attitude; details depend on the account. Under some accounts, indeterministic reasoning departs strongly from Broomean reasoning, by including automatic processes or reasoning from absences.

But on some versions of the first, third and fourth account, indeterministic reasoning share an important feature with Broomean reasoning: it is an activity, i.e., something

[16] While you can derive the 'train' intention from the four premise-attitudes (including the introspective belief) you cannot derive it after adding the 'boat' intention as a fifth premise-attitude. Such reasoning would rest on a strange combination of premise-attitudes, as the 'boat' intention (the fifth premise-attitude) clashes with the belief of not intending either means (fourth premise-attitude). Like under (i)–(ii) in the third account, you either do not even start reasoning to the 'train' intention, because your 'boat' intention is conscious. Or you start reasoning, but then, after bringing to mind some or all of the premise-attitudes, your remember your 'boat' intention, which lets you stop reasoning as you cannot derive the 'train' intention with your 'boat' intention as an additional premise-attitude.

you *do*, as opposed to something happening partly automatically (where this activity can be done explicitly, like Broomean reasoning). We now explain why, using the Venice example. We must ask two questions:

- First, is a conclusion-attitude – the 'boat' or 'train' intention – formed *actively*? On the first account, you first reason to a broad intention, which is then refined; this two-stage process is an activity provided the refinement happens actively rather than automatically. On the third account, you derive an intention directly by Broomean reasoning, hence actively. On the fourth account, the process of creating an intention is active provided that the additional premise-belief in your reasoning, i.e., the belief of not yet intending a means, is (if not already present) created actively by introspection rather than automatically. We discuss introspection in Section 3.4.
- Second, is an additional conclusion-attitude avoided *actively*? On the first account, it is avoided actively or automatically, depending on whether the conclusion-refinement process that produces a conclusion and avoids 'overproduction' is active or automatic. On the third account, surplus reasoning gets blocked by remembering the existing conclusion, which is an active process provided that you remember the intention actively (by calling it to mind) rather than automatically. On the fourth account, surplus reasoning gets blocked *either* because the already derived 'boat' intention lets you automatically lose your (by now false) introspective belief of holding neither intention, *or* because your 'boat' intention is conscious, *or* because you call it to mind actively, *or* because it comes to mind automatically. In the second and third case, surplus reasoning is avoided actively.

3.4 The mental activity of introspecting

What is introspecting? This question matters to the fourth account of reasoning. Introspecting is a mental activity that generates a belief about your attitudes, for instance the belief of not having certain attitudes. Such beliefs often appear automatically; but here we focus on the mental act of introspecting, by which you create such beliefs actively. Introspecting resembles observing (and thereby coming to believe) a fact, except that the fact in question is one about your mind. Introspecting is thus a mind-internal form of observing, which requires no external perception or sensory data.

Introspecting is an activity very different from reasoning, as it is not inferential. You do not *infer* something about your attitudes; you *observe* it. For instance, introspectively discovering your 'boat' intention differs from deriving the belief of having this intention. The categorical difference between introspecting and Broomean reasoning becomes even clearer when what you observe is the absence of attitudes. A belief of not having certain attitudes cannot be inferred from attitudes, but only reached by introspecting. Accordingly, in the fourth account of indeterministic reasoning, what makes you discover your lack of an intention is not reasoning, but introspecting.

While Broomean reasoning starts from premise-attitudes, introspecting arguably starts from a special attitude, namely one of *wondering (whether)*. In English, the term 'wondering' stands either for an attitude (as in 'I wonder what is going on') or

for an activity similar to searching (as in 'Despite wondering for a while what is going on, I did not find it out'). We always refer to the attitude. We leave open whether this attitude is a primitive attitude or a reducible attitude such as *intending to know whether* or *desiring to know whether*. In introspecting, you first wonder whether p (where p is a given proposition about your attitudes), which makes you look inside yourself and thereby observe whether p, which makes you believe p or believe *not* p.

Just as you can reason explicitly, you can introspect explicitly, by using the interrogative mood to express your wondering, as explained in Appendix A. So your indeterministic reasoning under the fourth account can be explicit in its entirety, including in the step of introspecting.

3.5 Indeterministic reasoning formalised

A full psychological model of reasoning facing choices would inevitably look different for the four accounts. We take a different strategy: we only model what is effectively achieved, namely the formation of a conclusion-attitude. Our 'shorthand model' thus fucuses on the result of reasoning on attitudes, while being silent on the psychological process. This makes the model account-neutral: it encompasses all four accounts, rather than committing to one.

The model generalises our deterministic model of Section 2.3 by including choices. It first generalises deterministic rules (Definition 2) into indeterministic rules:

Definition 4 *An **indeterministic reasoning rule** is a pair (P, K) of a set of premise-attitudes $P \subseteq M$ and a non-empty set of possible conclusion-attitudes $K \subseteq M$. A **(possible) revision** of a constitution C by this rule is a constitution given by*

$$C^+ = \begin{cases} C \cup \{k\} \text{ for some } k \in K & \text{if } P \subseteq C \text{ \& } K \cap C = \varnothing \text{ (the rule 'applies')} \\ C & \text{otherwise (the rule does not 'apply').} \end{cases}$$

An indeterministic rule with a single conclusion-attitude $(P, K) = (P, \{k\})$ is effectively deterministic, and can be identified with the corresponding deterministic rule (P, k). For there is a single revision of a constitution C by $(P, \{k\})$, namely the revision of C by the deterministic rule (P, k).

As in the deterministic model, you have many reasoning rules at your disposal – 'your' rules – and you reason with them to reach new constitutions. This leads to an indeterministic generalisation of Definition 3:

Definition 5 *An **indeterministic reasoning system** is a set S of indeterministic reasoning rules ('your' rules). Given an initial constitution C, a constitution C^+ is **reachable from C by (S-)reasoning** if there are a number of steps $n \geq 0$ and constitutions $C_0, .., C_n$ such that $C_0 = C$, $C_n = C^+$, and each C_t with $t \neq 0$ is a revision of C_{t-1} by some rule in S.*

An indeterministic reasoning system S containing only deterministic rules $(P, K) = (P, \{k\})$ can be identified with a deterministic reasoning system, by identifying each rule $(P, \{k\})$ in S with (P, k).

An indeterministic rule (P, K) is a metaphor or shorthand for a complex psychological process, whose nature depends on the account. None of the four accounts we presented *literally* involves an indeterministic rule. The first, third and fourth account involve ordinary deterministic rules, and the second account involves non-ordinary deterministic rules (with attitude-*absences* as additional premises). A full psychological model would therefore look rather different from our shorthand model. While the shorthand model ascribes to you an indeterministic rule (P, K), a full psychological model on one of the accounts invokes for each possible conclusion-attitude k in K a deterministic rule that derives k based on certain premises;[17] and it invokes a mechanism for 'handling' deterministic rules such as to effectively mimic indeterministic reasoning. Appendix C sketches a full psychological model on each account.

3.6 The structural properties of reasoning in attitudes

Reasoning in attitudes, using a given reasoning system S, defines a *consequence operator*, in fact a consequence operator of a special 'rule-based' type. In general, a consequence operator over the set M transforms old into new constitutions (subsets of M). Consequence operators are familiar objects in logic ever since Tarski (1956), although they usually operate over propositions or sentences, not over attitudes towards them. This interpretive difference does not affect the abstract analysis of consequence operators.

This section defines our 'rule-based' consequence operators formally, and investigates their structural properties. As we shall see, they are in one sense entirely standard: they are monotonic, inclusive, and idempotent. In another sense, they are non-standard: they are indeterministic, i.e., multi-valued, provided some rules in S are indeterministic.

A **deterministic consequence operator (over M)** is a function transforming each constitution $C \subseteq M$ into a new constitution $C^+ \subseteq M$, the *consequence or revision* of C. More generally, an **indeterministic consequence operator (over M)** is a function transforming each constitution $C \subseteq M$ into a set of constitutions, the *(possible) consequences or revisions* of C. If the operator transforms each constitution C into a singleton set $\{C^+\}$, i.e., if each C has a single revision, then we treat the operator as *deterministic*, by identifying an output set $\{C^+\}$ with its single element C^+.[18]

Your reasoning system S – whether deterministic or indeterministic – yields a consequence operator, called 'rule-based'. How is it defined? Recall that your initial constitution C evolves as you reason with rules in S. A constitution C^+ is reachable from C by (S-)reasoning if it emerges after applying any number $n \geq 0$ of rules from S in any order (see Definitions 3 and 5). If moreover no rule from S applies to C^+, i.e., no further reasoning is possible, then we call C^+ 'reachable by *maximal* (S-)reasoning from C'. Maximal reasoning yields our rule-based operator:

[17] On the first and third account the premise-attitudes are the attitudes in P; on the second account they are the attitudes in P *plus* the absences of the attitudes in K; on the fourth account they are the attitudes in P *plus* the belief of not yet possessing attitudes in K.

[18] Some logicians study multi-conclusion inference (e.g., Restall, 2005; Beall, 2011). Multi-conclusion inference resembles our reasoning with choices in that it also generates an indeterministic reasoning operator, but often differs in that it only requires to draw at least one conclusion, not necessarily exactly one.

Definition 6 *The **consequence operator given by a reasoning system** S is the consequence operator Cn_S that transforms any constitution C into*

$$Cn_S(C) = \{C^+ : C^+ \text{ is reachable from } C \text{ by maximal } S\text{-reasoning}\}.$$

*A consequence operator is **rule-based** if it is given by some reasoning system.*

Defining rule-based operators by *maximal* reasoning implicitly assumes that all your rules *must* be applied, i.e., that you face no choice-whether in reasoning, only a choice-between (see Section 3.1 for this distinction). Choice-whether is neglected in Definition 6 because it is not our primary focus. But one can generalise the notion of rule-based operators and the below theorem by permitting choice-whether.[19]

Rule-based operators Cn_S are usually indeterministic, but becomes deterministic if all reasoning rules in S are deterministic. One might call rule-based operators 'Broomean', but this would be a stretch if S contains indeterministic rules, which are not considered by Broome.

Deterministic consequence operators are far more common than indeterministic ones. Here are the three most standard properties of a deterministic operator $C \mapsto C^+$:

- *Inclusiveness*: For any initial constitution $C \subseteq M$, $C \subseteq C^+$ ('revision never removes attitudes').
- *Idempotence*: For any initial constitution $C \subseteq M$, $(C^+)^+ = C^+$ ('revising again has no effect').
- *Monotonicity*: For any initial constitutions $C, D \subseteq M$, if $C \subseteq D$ then $C^+ \subseteq D^+$ ('the more attitudes you have ex-ante, the more you have ex-post').

We generalise these properties to an indeterministic consequence operator $C \mapsto Cn(C)$:

- *Inclusiveness*: For any initial constitution $C \subseteq M$ and revision of it $C^+ \in Cn(C)$, $C \subseteq C^+$.
- *Idempotence*: For any initial constitution $C \subseteq M$ and revision of it $C^+ \in Cn(C)$, the only revision of C^+ is C^+, i.e., $Cn(C^+) = \{C^+\}$.
- *Monotonicity*: For any initial constitutions $C, D \subseteq M$, if $C \subseteq D$ then each revision $D^+ \in Cn(D)$ satisfies $C^+ \subseteq D^+$ for some revision $C^+ \in Cn(C)$.

The triple of indeterministic definitions reduces to the standard triple if the consequence operator is deterministic.

Our rule-based consequence operator has all three properties, and is in this sense classical, albeit indeterministic.

Theorem 1 *Any rule-based consequence operator (given by a deterministic or indeterministic reasoning system) is inclusive, idempotent, and monotonic.*

[19] How? Partition S into a set S_1 of 'mandatory' rules (which you *must* apply, leaving no choice-whether) and a set S_2 of 'optional' rules (which you *can* apply, leaving a choice-whether). S_1 represents your reasoning obligations, S_2 your reasoning permissions. The generalised rule-based operator Cn_{S_1,S_2} is definable like Cn_S, merely replacing 'maximal S-reasoning' with 'S-reasoning that is maximal w.r.t. S_1' (suitably defined). Cn_{S_1,S_2} reduces to Cn_S if $S_1 = S$ and $S_2 = \varnothing$, i.e., if all rules are mandatory. Theorem 1 generalises *partly* to Cn_{S_1,S_2}, which remains inclusive and monotonic, but can become non-idempotent.

Part of this result is obvious. Rule-based consequence is inclusive, as reasoning never removes attitudes; and it is idempotent, as no further change is possible after maximal reasoning (but this changes when adding choice-whether, by fn. 19). Monotonicity is less trivial; it is established in the appendix.

4 The difficulty to model reasoning in attitudes logically

Given that (deductive) theoretical reasoning follows logical entailment, one might expect that reasoning in attitudes also follows entailment of some sort, and hence could be modelled logically. By a logical model of reasoning we mean a model of reasoning as following some entailment relation.

We now go, largely unsuccessfully, through the three most natural attempts to model reasoning in attitudes logically (Section 4.1–4.3). To rule out trivial deviations from entailment, much of the discussion will implicitly assume that you are an 'ideal' reasoner, i.e., that your reasoning system S contains the right kind of rules. We do not spell out what this means exactly.[20]

4.1 Content entailment: a model of reasoning in a *single* attitude

When theoretical reasoning is said to follow entailment, one refers to entailments between belief-*contents*. So, our first attempt must be to model reasoning in attitudes by entailments between attitude-*contents*.

Reasoning in attitudes would have no chance to follow entailment between attitude-contents if those contents were not propositions or something similar. We would commit a category error by invoking entailments between the contents of attitudes like *adoring Paris* and *imagining a fireball*, i.e., entailments between things like towns and fireballs. The 'propositionalist' view that all intentional attitudes have, or can be reduced to having, propositional content is controversial (Montagne 2007). While not committing to full-blown propositionalism, we do assume here that *reasoning* in attitudes is reasoning in *propositional* attitudes.

So again: does reasoning in (propositional) attitudes follow entailment between the propositions? Logic is in the first place about propositions, i.e., the world, not about attitudes. But insofar as propositions are the contents of attitudes, logicians can indirectly address attitudes. When logicians do so, they notoriously choose *beliefs*: they interpret propositions as *belief*-contents, which turns logical entailment into a model of reasoning in beliefs, not in desires, or in intentions, etc. Could logicians instead choose desires (or intentions, etc.), and take content entailment to model reasoning in desires (or intentions, etc.)? Such a model would support reasoning from desiring p into desiring p *or* q (or from intending p into intending p *or* q, etc.), as p entails p *or* q; and it would

[20] Intuitively, S contains only good rules ('soundness') and sufficiently many rules ('completeness'). Examples of good rules might be the deterministic rules in (a)–(d) in Section 2.2 and the indeterministic rule in the Venice example. To be complete, S might need to contain these and many other rules; for instance, without the rule in the Venice example you are handicapped, i.e., unable to form a 'boat' or 'train' intention where needed. For deterministic rules, being good might mean being correct in Broome's sense.

support reasoning from nothing into desiring a tautology (or intending it, etc.), as the empty set entails the tautology. One might doubt such reasoning, and hence reject the idea that reasoning in desires (or in intentions, etc.) follows content entailment.

One reason for this doubt is that entailment is deterministic, whereas reasoning in desires (or in intentions, etc.) might involve choices. We see no principled reason to exclude such choices, although convincing examples seem harder to find than when mixing attitude types, as done in the Venice example.

But even if reasoning in desires (or in intentions, etc.) followed content entailment, we would not have modelled general reasoning in attitudes. Reasoning in desires (or in intentions, etc.) is still mono-attitude reasoning. Once we mix attitude types, as practical reasoning routinely does, content entailment obviously cannot model reasoning: while p and *if p then q* entail q, you would not reason from hating p and believing *if p then q* into intending q.

In sum, although content entailment can famously model theoretical reasoning (with qualifications discussed in Section 5), it is debatable as a model of reasoning in a given non-belief attitude such as desire or intention, and fails for reasoning in multi-attitudes.

For even simpler reasons, reasoning in *non-monadic* attitudes cannot follow content entailment, as such attitudes have complex contents. For instance, reasoning in preferences (Broome 2006) is reasoning in attitudes towards *pairs* of propositions. Entailments go between propositions, not between pairs.

4.2 Attitude entailment: a model of reasoning *about* attitudes

The mismatch between reasoning in attitudes and content entailment already shows that any relation between reasoning in attitudes and logic – if existent – is very different from that for theoretical reasoning. Such a different relation would exist if reasoning in attitudes followed *attitude* entailment. However, read literally, attitude entailment models theoretical reasoning *about* attitudes, not reasoning *in* attitudes. Why? Assume you reason in attitudes as follows:

I ought to pay taxes. So, I shall pay taxes. (5)

Here you reason from a belief to an intention, following an instance of the Enkratic Rule in Section 2. Does this reasoning follow an attitude entailment, namely (in formal-logical terms) the entailment $B(p) \models I(q)$, where B is a belief operator, I is an intention operator, and p and q are sentences representing *I ought to pay taxes* and *I will pay taxes*, respectively? Whenever we state attitude entailments in formal logic, we presuppose a suitable logic of attitudes, with modal operators for all relevant attitude-types, e.g., a belief operator, an intention operator, or a (dyadic) preference operator. (Logics of attitudes exist in abundance. They can do many things.[21])

[21] Mono-modal logics address one attitude, e.g., belief in 'doxastic logics' (e.g., Halpern 2017) and preferences in 'preference logics' (e.g., Liu 2011). Multi-modal logics address more than one attitude, e.g., belief, desire and intention in 'BDI logics'. Logics of attitudes capture rationality of attitudes by axioms (e.g., axioms requiring that tautologies are believed).

Yet the literal reading of the entailment $B(p) \models I(q)$ is not (5), but this:

I believe I ought to pay taxes. So, I intend to pay taxes. (6)

Here you reason *about* your attitudes: you deduce you have an intention from having a belief. In (5) you do not reason about your attitudes: you are not your own observer who recalls having a belief and deduces ('discovers') an intention, but you *form* an intention. Attitude entailment models attitude discovery, not attitude formation. It models reasoning about, not in attitudes. Reasoning about attitudes does not change *these* attitudes, but it creates (meta-)beliefs about them (cf. Broome 2014 and Dietrich et al. 2019).

Worse, the different reasoning (6) which the entailment $B(p) \models I(q)$ models is invalid as an inference about your attitudes: its premise can hold without its conclusion holding. Indeed, before starting your reasoning (5), you believed your ought to pay taxes without (yet) intending it; formally, $B(p)$ was true and $I(q)$ was false. Why, then, does the logic deem the inference $B(p) \models I(q)$ valid? Nothing is wrong with the logic, but we have misapplied it. The logic is one of *rational* attitudes, unlike your imperfect attitudes. Attitude entailment in this logic represents reasoning about *rational* attitudes.

The point of reasoning *in* attitudes is to become more rational. Ironically, reasoning *about* your attitudes works if your attitudes are already rational (or more generally display regularities, as explained below), whereas reasoning *in* attitudes matters if your attitudes are not yet rational.

To be precise, improving rationality need not be the purpose of reasoning in attitudes. You could for instance reason as in Section 2:

Paying taxes is legally required. So, I shall pay taxes. (7)

This reasoning from a belief to an intention is not enkratic reasoning, because it starts from a belief about what is *legally* required, not what you ought to do. As Broome might say, you reason towards legality, not rationality. Modelling (7) by an entailment – namely by $B(p') \models I(q)$ where p' represents the new premise content – is again problematic, still because the entailment represents a piece of reasoning *about* attitudes.[22] The novelty of the example is that the entailment $B(p') \models I(q)$ is invalid in a logic of 'merely' rational attitudes, as the premise-belief does not *rationally* entail the conclusion-intention. The entailment holds in a logic of ideal attitudes in a suitably comprehensive sense of 'ideal' that captures the relevant norms, such as rational, legal, or moral norms.

In sum, attitude entailment in a suitable logic of rational (or otherwise ideal) attitudes does not model reasoning in attitudes, but reasoning about attitudes, more precisely about rational (or otherwise ideal) attitudes. So, reasoning about your attitudes presupposes that they are rational (or otherwise ideal).

This said, in principle you can also reason about irrational or otherwise non-ideal attitudes (of yourself or someone else), as long as they display *some* systematic patterns or regularities, possibly irrational patterns. You might reason as follows about Ann who always intends the opposite of what she believes she ought to do:

Ann believes she ought to pay taxes. So, she intends not to pay them.

[22] The entailment reads: *I believe paying taxes is legally required. So, I intend to pay taxes.*

But you cannot reason about 'chaotically irrational' attitudes, because there are no non-trivial entailments between them.

4.3 Attitude entailment: an as-if model of reasoning in attitudes?

Could attitude entailment at least serve as an as-if model of reasoning in attitudes, instead of a literal model? That is, could reasoning in attitudes change your attitudes in a way that *effectively* mimics attitude entailment, so that attitude entailment predicts the effect of reasoning? This as-if approach is psychologically silent: it does not model the psychological process (the passage from premises to conclusions), but only the resulting attitudinal changes. Reasoning would then behave as if following attitude entailment, hence be extensionally equivalent to attitude entailment.

This section's shift in focus from the internal psychology of reasoning to the external effect of reasoning on attitudes may seem like a shift of project. However, just as the notorious question of whether rational-choice theory can appropriately model decision-making has psychological and behavioural dimensions, so our own question of whether logic can appropriately model reasoning in attitudes has both dimensions. Going for an as-if interpretation is a classic maneuver used to defend rational-choice models.[23] Although this maneuver is increasingly controversial (cf. Hausman 2012, Dietrich and List 2016, and Guala 2019), we should give the as-if approach a fair chance in the context of reasoning. This is why we now ask whether logical entailment can yield an as-if model of reasoning in attitudes.

Such an as-if model must however fail, because reasoning is an indeterministic process whereas attitude entailment is deterministic. We now spell this fact out precisely, by formulating and later rejecting the hypothesis of extensional equivalence between reasoning and attitude entailment. This is the equivalence hypothesis on which the as-if model rests:

Extensional Equivalence Hypothesis (EE) – informal statement: You reason to an attitude if and only if your initial attitudes entail that attitude.

The hypothesis can be re-stated formally, after formalising reasoning and entailment:

- Your reasoning is based on your reasoning system S, and is formally captured by the consequence operator Cn_S that transforms each constitution $C \subseteq M$ into a (usually non-unique) revision, i.e., into any constitution C^+ reachable from C by maximal S-reasoning (cf. Section 3.6).
- Attitude entailments are formally given by the entailment relation of a suitable logic of (rational or otherwise ideal) attitudes. Each attitude $m \in M$ is represented by a sentence saying that you have this attitude; it is denoted m^* and takes the form $O(\phi)$ where O is the relevant attitude operator and ϕ is the relevant sentence. For instance, if m is *intending to swim*, then m^* is $O(\phi)$ where O is the intention

[23] Under an as-if interpretation, a standard rational agent behaves as if maximising expected utility, but utilities and probabilities carry no psychological meaning, merely representing behaviour (cf. Cozic and Hill 2015). Under literal or mentalist interpretations, utilities and probabilities are psychological constructs capturing values and beliefs.

operator and ϕ reads 'you swim'.[24] An entailment from your initial attitudes (in C) towards a new attitude k is thus formalised as $\{m^* : m \in C\} \models k^*$. In words: the attitudes in C (rationally or otherwise ideally) entail the attitude k.

Extensional Equivalence Hypothesis (EE) – formal statement: For any initial constitution $C \subseteq M$, revision of it $C^+ \in Cn_S(C)$, and attitude $k \in M$,

$$k \in C^+ \Leftrightarrow \{m^* : m \in C\} \models k^*. \tag{8}$$

The left side of the equivalence (8) says that after reasoning maximally you have the attitude k. The right side says that k is entailed by your initial attitudes. In the example (6), k is *intending to pay taxes*, logically represented as $I(q)$ $(= k^*)$, and your initial constitution C contains *believing you ought to pay taxes*, logically represented as $B(p)$. Here EE says: this intention is among the attitudes formed by reasoning if and only if it is entailed by your initial attitudes, formally, $k \in C^+ \Leftrightarrow \{m^* : m \in C\} \models I(q)$.

Such an as-if model of reasoning is psychologically silent. It captures reasoning to k by an abstract condition ('$k \in C^+$') that tells only that you effectively form attitude k, without revealing the procedural fact of how many reasoning steps are needed to derive k from C and which initial attitudes from C enter your reasoning as premise-attitudes; you might reach k in just one step from just one premise (as in your reasoning in (6)) or in many steps drawing on many initial attitudes. Since the premises of reasoning are left out, your reasoning is not modelled by an entailment from your premise-attitudes, but by an entailment from the totality C of initial attitudes (formally, by $\{m^* : m \in C\} \models k^*$).

In sum, the as-if model based on EE treats reasoning as a black box that produces new attitudes by entailment from initial attitudes, regardless of the psychological mechanism at work. This procedural blindness reflects the reduced ambition of the as-if approach, which aims to model what reasoning achieves *in effect*, not *how* it achieves it – an approach we took reluctantly after the more substantive and mentalistic attempts had failed.

But EE is not tenable, and with it the as-if approach. We first discuss the central objection, based on choice in reasoning and directed against the necessity of entailment for reasoning (direction '$\Rightarrow$'). We then mention an objection against sufficiency (direction '$\Leftarrow$'), and an ad-hoc-ness concern.

Against necessity. You can face choice in reasoning – either choice-whether (reason or not reason?) or choice-between (reason to which conclusion?). If you face choice and reason to an attitude, then that attitude is not entailed by your initial attitudes, as you could have abstained from reasoning or reasoned otherwise.

As usual (and as in our formal statement of EE[25]), we focus on choice-between. In our Venice example, you reason to a 'boat' intention, which is not entailed by your

[24] Presumably, the assignment $m \mapsto m^*$ defines a bijective correspondence between M and the set of logical sentences of type $O(\phi)$ for some attitude operator O.

[25] Our formal statement of EE excludes choice-whether because the reasoning operator Cn_S excludes it (being defined by *maximal* reasoning). A version of EE that includes choice-whether is obtained by replacing Cn_S with the generalised operator defined in fn. 19. This version would be more clearly false.

initial attitudes, as you could form the 'train' intention instead. (Recall: 'entails' means 'rationally entails', or more generally 'ideally entails' under some sense of ideal attitudes.)

If anything, a modified version of EE holds, whereby entailment determines which attitudes you *necessarily* reason to, i.e., belong to *all* revisions C^+ of your initial constitution C.[26]

One might defend EE against the 'choice' objection by rejecting the very idea of choice in reasoning. One might instead make two claims: (i) you never reason to a 'boat' intention, but to the broad intention to take a boat *or* a train, this intention being later refined by some other process; (ii) the broad intention *is* rationally entailed by your initial attitudes. We doubt both claims. Claim (i) follows the first account of choice in reasoning, which is actually an account of *deterministic* reasoning followed by conclusion-refinement (cf. Section 3.2). No doubt, often in life you reason to broad intentions that get refined later. Yet you also often reason directly to a specific intention, facing a choice between possible intentions. So the 'never' in Claim 1 is false.

Claim (ii) initially seems correct. It seems that your intention to visit Venice and your beliefs about possible means rationally entail intending to take a boat or train, following the requirement (*) 'intend what you believe is a necessary means to your intended end'. Perhaps Claim (ii) is true. We however see two counter-arguments:

- On one counter-argument, taking a boat or a train is not a means, but a disjunction of means. Means, on this argument, are special things, presumably certain actions or certain causes (in a 'production' sense rather than 'difference-making' sense). Taking a boat and taking a train are two actions or causes of relevant kind, but their disjunction is not, because it is somehow non-unified. What you need to intend according to (*) is a necessary *means*, not disjunction of means. So your initial attitudes do not rationally entail intending the disjunction of means. Indeed, you can perfectly reach Venice without ever forming this broad intention, as long as you intend a means, e.g., intend to take a 'boat'. You may form the broad intention as an intermediate step towards intending a means, but this step is not required.[27] This argument rests on a sophisticated notion of a means that is tricky to spell out precisely. A difficulty is that any action can be refined. For instance, taking a boat is the disjunction of taking a large or small boat. Why then should taking a boat qualify as a means, while taking a boat or train does not? Presumably because taking a boat is something unified or connected (albeit refinable), while taking a boat or train is something disunified or disconnected, hence not an action or cause of relevant type.[28] Borderline cases come immediately to mind. If taking the left

[26] The modified hypothesis is: You necessarily reason to an attitude if and only if your initial attitudes entail that attitude. Formally, for any initial constitution $C \subseteq M$ and attitude $k \in M$,

$$k \in \cap_{C^+ \in Cn_S(C)} C^+ \Leftrightarrow \{m^* : m \in C\} \models k^*.$$

[27] In this context, note that you may intend something without intending its implications; e.g., you may intend to take a boat without intending to take a boat or a train.

[28] To draw an analogy with mathematical topology: in the space of logically possible actions, the set of actions of taking some boat is topologically connected, unlike the set of actions of taking some boat or some train.

door and taking the right door are two means of getting outside, is then taking *some* door insufficiently unified for being a means, or is it just unified enough for being a means? We thank a referee for raising this issue. Much hinges on the notion of 'unified', which might come in degrees and contain vagueness, and which could be determined on mainly semantic grounds or mainly physical grounds. The next counter-argument avoids these complications, by being liberal about the notion of means.

- On another counter-argument, due essentially to Broome (2013), even if taking a boat or train *is* a means for visiting Venice, more precisely is believed to be a necessary means to that intended end, you are not required to intend it, because it can come about without being intended. Indeed, intending a specific means (e.g., to take a boat) ensures not just this specific means, but also the broad means. This argument amounts to rejecting the rationality requirement (*). One could replace it with a weaker requirement: 'if you believe something is a necessary means to your intended end, then intend something that you believe implies this means'. Or, following Broome (2013: 159), one could refine (*) by adding the condition that (you believe that) intending the means is necessary for the means. Both refinements of (*) block the undesired conclusion of having to intend to take a boat or train in the Venice example.

Under either counter-argument, your initial attitudes do not rationally entail the broad intention. They also do not rationally entail any given specific intention. They rationally require holding *some* specific intention, as a matter of instrumental rationality; cf. Broome's (2013: 170) 'generalised instrumental requirement'.

Against sufficiency. Sometimes you cannot form an attitude although your attitudes entail it. Perhaps you are akratic, and unable to form an intention k which rationally follows from your beliefs about what you ought to do. Here your constitution entails k (formally, $\{m^* : m \in C\} \models k^*$), but you cannot reason to k (formally, $k \notin C^+$ for every revision C^+ of C). Perhaps you believe that attitude k makes you happy; this belief (let us assume) rationally entails forming k, which you are unable to do. Perhaps you intend to become wise and believe wisdom requires studying, but you are psychologically unable to intend to study, although this intention is (rationally) entailed.

However, these counterexamples apply to an imperfect reasoner, who is unable to perform some correct reasoning. One can rehabilitate sufficiency by assuming a perfect reasoner who does not suffer from psychological 'reasoning barriers'. This reasoner's reasoning system S not only contains only correct rules, but also contains sufficiently many rules, enabling barrier-free reasoning towards rationality. This perfect reasoner might satisfy the sufficiency claim, and the modified hypothesis in fn. 26.

Ad-hoc-ness charge: Attitude entailments are entailments between attitude propositions of the simplest type: propositions saying that you possess a certain attitude, e.g., that you desire p. Call them *atomic* attitude propositions. There exist many *non-atomic* attitude propositions: that you do *not* desire p, that you desire p *and* believe q, etc. Entailments between non-atomic attitude propositions do not correspond to reasoning

in attitudes. For instance, the entailment $\{B(p) \vee I(q), \neg D(r)\} \models \neg D(s)$ (for operators of belief B, intention I, and desire D) does not correspond to any reasoning in attitudes, because such reasoning cannot start from or generate disjunctions or absences of attitudes. You can reason *about* absences or disjunctions, but not *in* them (cf. Dietrich et al. 2019). It might seems ad hoc to pick out particular entailments – those between *atomic* attitude propositions – and grant them a perfect correspondence to reasoning, while denying such a correspondence for all other entailments.

5 Concluding remarks: reasoning in attitudes vs. theoretical reasoning

Where do we stand? Reasoning in attitudes differs fundamentally from reasoning about attitudes, and more generally from theoretical reasoning. Conceptual differences aside, reasoning in attitudes does not follow entailment between attitude-contents – unlike (deductive) theoretical reasoning. It does also not follow entailment between attitudes – unlike (deductive) reasoning about your attitudes. The main reason is choice in reasoning. There may be choice between rival conclusions (choice-between) and choice whether to derive a conclusion at all (choice-whether). We have mainly addressed choice-between. Our focus lied on how you avoid surplus conclusions (question (ii) in Section 3.1), not how you reach any conclusion at all (question (i) in Section 3.1). This was possible by adopting Broome's (2013) position that standard (Broomean) reasoning already solves the second problem, i.e., allows you to derive a conclusion even when you face a choice. This position is interesting. It is bound to be controversial, as it essentially implies that Buridan's ass faces no real dilemma, i.e., can reach one of the stacks of hey by the same simple reasoning by which it could reach it without facing a choice. If one disagrees, then one could argue that reasoning with choice becomes possible by using a *tie-breaking rule* that selects a conclusion from the given set of possible conclusions. There are many possible tie-breaking rules, such as to choose like you did last time you faced an analogous choice. How and whether tie-breaking can enter your reasoning is a problem of its own.[29]

The special status of theoretical reasoning might surprise. After all, such reasoning is simply a special kind of reasoning in attitudes, namely reasoning in beliefs. Why then can (deductive) theoretical reasoning follow entailment? The central point is the absence of choice in theoretical reasoning: you do not choose between different possible conclusions when forming beliefs rather than, say, intentions. For beliefs track an external truth. They aim to match the external world. As the external world obeys logic, so does (deductive) theoretical reasoning.

[29] Under one approach, tie-breaking is an explicit part of reasoning that requires additional premise-beliefs, for instance the belief that such-and-such possible conclusion is selected by such-and-such tie-breaking rule, *plus* the intention to use this rule. Here your reasoning is only apparently indeterministic: you actually reason deterministically, with some implicit premises that break a tie. Yet the problem of indeterminacy might reemerge, as you need to come to intend a tie-breaking rule, thereby facing a choice between tie-breaking rules. Under another approach, tie-breaking rules intervene automatically. Might this undermine the idea of reasoning as a mental act? We cannot settle these and similar issues here.

There are well-known exceptions, where theoretical reasoning arguably departs from content entailment. You might derive *more* beliefs than are entailed, by reasoning inductively. You might derive *fewer* beliefs than are entailed, because subjectively probable (believed) propositions sometimes jointly entail subjectively improbable (disbelieved) propositions, as in the Lottery Paradox (Kyburg 1961, Dietrich and List 2021). We say 'might' because our Broome-inspired account of reasoning might escape at least the second phenomenon, as explicit theoretical reasoning might exclude implicit probabilistic considerations.[30] There might be a third (controversial) reason for departure from entailment: non-epistemic motives, such as the pursuit of happiness. It is certainly questionable whether you can form a belief in order to become happy, even if you wanted to. But if it did work, then it would strongly disconnect theoretical reasoning from content entailment, perhaps to the point of introducing choice into theoretical reasoning.

Still, content entailment is a first-order approximation of theoretical reasoning. By contrast, reasoning in general attitudes goes beyond logic. No doubt, it can be studied formally, namely by using indeterministic consequence operators, as we have started to do. But such tools are 'logical' only in an abstract sense.

Appendix

A Explicit introspection

While Broomean reasoning is a process of bringing to mind premise-attitudes and then creating a new attitude, introspecting is arguably a process of bringing to mind your wondering and then creating a new meta-belief – although this belief is not *derived* from or *based on* your wondering, as introspecting is not an inferential process, unlike reasoning. Like Broomean reasoning, introspecting can in principle be done explicitly. How? By saying to yourself the marked content of your wondering and your resulting belief. Using the interrogative mood as a linguistic marker of wondering, you may introspect explicitly as follows in the fourth account of indeterministic reasoning:

Do I already intend to take a boat or train? I do not yet intend either.

You say no 'So', as you draw no inference, unlike in Broomean reasoning.

This suggests an objection against the fourth account: if indeterministic reasoning is indeed premised on a meta-belief whose formation by introspection requires wondering (as just claimed), then you cannot reason without wondering in the first place. In the Venice example: if reasoning to a 'boat' or 'train' intention requires finding out introspectively that you have neither intention, which requires wondering whether you do, then you cannot reason without initially wondering about this. But normally you do not initially wonder about this (why should you?). Technically, if your initial constitution

[30] You may reason in *partial beliefs* (cf. Staffel 2013); and you can do so explicitly, using markers such as 'probably'. Such reasoning does of course not follow entailment between the contents of partial beliefs. But by 'reasoning in beliefs' we mean 'reasoning in straight beliefs'.

C does not contain the relevant attitude of wondering, you cannot introspect, hence cannot reason.

What could help? You might be lucky and start wondering automatically when needed. But if the fourth account relied on 'automatic wondering', we could not uphold our claim that indeterministic reasoning on this account can be a mental activity (cf. Section 3.3).

Surprisingly, however, you can come to wonder about something by an act of Broomean reasoning. From what premises do you derive a wondering? Intuitively, the premises are the justifications or basis for wondering. In the Venice example, your wondering whether you already have a 'boat' or 'train' intention might be derived from your intention to visit Venice and your beliefs about possible means. Then the complete process by which you form a 'boat' intention (on the fourth account) is made explicit by the following inner speech:

(1) I shall visit Venice. (2) For this I must either take a boat or take a train. (3) Both means are equally good. (4) Do I already intend a means? (5) I do not intend a means. (6) So, I shall take a boat. (9)

In (1)–(3) you bring to mind attitudes; from (1)–(3) you derive the wondering (4) by reasoning; from (4) you reach (5) by introspecting; from (1)–(3) and (5) you derive (6) by reasoning.[31]

Our claim that you can derive a wondering by reasoning is certainly debatable. The claim is perhaps easier to accept if, as suggested in Section 3.4, *wondering whether* is a derived attitude, such as *intending (or desiring) to know whether*. In this case, reasoning to wondering is reasoning to a particular intention or desire. In our Venice example it seems very natural that, based on your initial attitudes, you derive an intention or desire to know whether you already hold specific intentions as to how to reach Venice. This derivation goes so quickly and easily that it rarely happens explicitly.

B Correct indeterministic reasoning?

According to Broome (2013), you can reason *correctly* to one of several possible conclusions. That is, given suitable sets P of premise-attitudes and K of possible conclusion-attitudes, you can reason correctly from P to any attitude in K. Broome focuses on instrumental reasoning, in which K contains intentions of a means to an end; for instance, in our Venice example K contains the 'boat' and 'train' intention, and P contains the intention to visit Venice and certain beliefs (about which Broome is more sophisticated than us, as mentioned in fn. 15). We have questioned Broome's correctness claim, suggesting that correctness requires adding a premise-attitude to P, namely the belief of not yet having any attitude from K.

We here discuss the issue. For argument's sake, we adopt Broome's general characterisation of correctness: reasoning from a set of premise-attitudes P to a conclusion-

[31] On an alternative view, your wondering is derived from *no* premises, as wondering needs no grounds. Then your complete inner speech might again take the form (9), but this time your wondering in (4) is derived from no premises rather than from (1)–(3).

attitude k is *correct* if P is a rationally permissible basis of k, more precisely, if it is rationality permitted to (i) hold the attitudes p in P at some moments and (ii) hold the attitude k at some moment *based on* the former attitudes. This so-called 'basing permission' is diachronic: it relates your attitudes at different moments. Indeed, reasoning takes time. By the time you reach the conclusion, you may have lost some premise-attitude(s), despite the conclusion-attitude being based on the premise-attitudes. Broome does not analyse what 'based on' means; nor shall we.

Broome applies this correctness criterion to certain examples of reasoning with different possible conclusions, similar to the Venice example. He claims that reasoning to any possible conclusion passes the correctness test. Technically, if P and K are the relevant attitude sets, say those in the Venice example, then for any k in K the rule (P, k) passes the correctness test, i.e., in short, P is a permissible basis of k. We doubt this claim. Were it true, you could reason correctly to all attitudes in K one by one, thereby forming conflicting intentions in the Venice example. Arguably, Broome has underspecified the basis of a k in K: a permissible basis is not P, but $P \cup \{m\}$, where m is the belief of not (yet) having any attitude from K. One could replace m by other introspective beliefs.[32] It is hard to say whether one could replace m by a non-introspective ('first-order') attitude.[33]

After deriving the conclusion-attitude k from $P \cup \{m\}$, your introspective belief m is false. So rationality does not permit holding the conclusion-attitude and all premise-attitudes in $P \cup \{m\}$ *simultaneously*. Yet rationality permits holding these attitudes at different times. This is why the reasoning rule $(P \cup \{m\}, k)$ can meet the correctness test.

C Full psychological models of the four accounts of choice in reasoning

The main text worked with a shorthand model of choice in reasoning, which is account-neutral thanks to focusing on the ultimate effect of reasoning on attitudes rather than the psychological process. This appendix sketches how a full psychological model might roughly look under each account. In fact, we only discuss how to model a particular instance of choice in reasoning: reasoning from a set of premise-attitudes $P \subseteq M$ to any attitude from a set of possible conclusion-attitudes $K \subseteq M$ (in our Venice example, P contains the intention to visit Venice and two beliefs, and K contains the 'boat' intention and the 'train' intention). The shorthand model represents this instance by the indeterministic rule (P, K). A full model might instead take the following form.[34]

[32] An alternative basis of k might be the set $P \cup \{m_l : l \in K \backslash \{k\}\}$, where m_l is the belief of not possessing attitude l.

[33] Is $P \cup \{m'\}$ a permissible basis of k if m' is, say, the intention not to take a boat or train to Venice, or the intention to travel to India (which prevents taking a train or boat to Venice)? This hinges on the notion of basis, which Broome and we find hard to spell out. Perhaps $P \cup \{m'\}$ is a permissible basis. If not, perhaps it becomes one after adding the belief that the attitude m' implies not intending to take a boat or train to Venice. But a so-increased basis would again contain an introspective belief.

[34] After knowing how to model an instance of indeterministic reasoning, one can easily model indeterministic reasoning in general. For the shorthand model, this step was taken by introducing a reasoning

First account. Here, a full model of reasoning from P to any attitude in K has two ingredients. One is a deterministic rule $(P, k^\vee)$ where $k^\vee$ is a suitable 'broad' or 'disjunctive' conclusion-attitude (in the Venice example: the 'boat or train' intention). The second ingredient represents the psychological process that refines your disjunctive attitude $k^\vee$ into an arbitrary attitude k in K (in the Venice example: into the 'boat' intention or 'train' intention). The composition of your reasoning rule $(P, k^\vee)$ and the automatic process is effectively equivalent to the indeterministic rule (P, K) in our shorthand model.

Second account. Here, a full model of reasoning from P to any attitude in K involves, for each possible conclusion attitude $k \in K$, a rule that derives k if all attitudes in P are present *and* all attitudes in K are absent; denote this rule by (P, K, k). These rules are generalised deterministic rules, premised on presences *and absences* of attitudes.[35] Reasoning with the rules (P, K, k) $(k \in K)$ is effectively equivalent to reasoning with the single indeterministic rule (P, K). Why? For any initial constitution C, either $K \cap C \neq \varnothing$, in which case none of these rules applies and the constitution stays C; or $K \cap C = \varnothing$, in which case any rule (P, K, k) applies and leads to the new constitution $C \cup \{k\}$, after which none of the rules applies anymore, so that the constitution stays $C \cup \{k\}$. The result is the same as for reasoning with the indeterministic rule (P, K).

Third account. Here, a full model of reasoning from P to any attitude in K involves, for each possible conclusion attitude $k \in K$, the standard deterministic rule (P, k), which forms attitude k if you have all attitudes in P. The model also contains a precondition for applying these rules: each of these rules can only be applied to constitutions not yet containing any attitude from K. This precondition operationalises the assumption that when you start reasoning from the premise-attitudes in P *but already possess* an attitude from K, then (on the account) your reasoning stops, caused by your having in mind or bringing to mind a preexisting attitude from K. The precondition prevents the rules from operating the usual way: they effectively operate like the generalised deterministic rules from our model of the second account (i.e., rules premised on the absence of attitudes from K). This is why reasoning in the current model is effectively equivalent to reasoning in the second model, and hence to reasoning in the shorthand model based on the indeterministic rule (P, K).

Fourth account. Here, a full model of reasoning from P to any attitude in K involves, for each possible conclusion attitude $k \in K$, the standard deterministic rule $(P \cup \{m\}, k)$, which forms attitude k from the attitudes in $P \cup \{m\}$, where the additional premise-attitude m is the belief of having no attitude from K (in the Venice example: the belief of having no 'boat' intention and no 'train' intention). We must also model the mechanism that prevents repeated reasoning to different attitudes from K (we shall only model

system S, and defining how constitutions can change by reasoning with S (Definition 5). For our four full models, one could proceed analogously, by introducing a construct analogous to a reasoning system and defining how constitutions can change by reasoning with this construct.

[35] A rule of this sort is definable as a triple (P_1, P_2, k), and adds the attitude k to a constitution C if $P_1 \subseteq C$ and $P_2 \cap C = \varnothing$.

the first such mechanism envisaged by the fourth account). To do this, we prescribe that each rule $(P \cup \{m\}, k)$ $(k \in K)$ applies in a non-standard way to a constitution C: systematically, *before application* the attitude m is added to your constitution, and *after application* m is removed again. Interpretation: while reasoning you get aware of not having any attitude from K, i.e., form the belief m, and after reasoning you lose that belief as it has become false through forming an attitude from K. More precisely, m is only added (and later removed) if $K \cap C = \varnothing$ and $P \subseteq C$. Why only then? If $K \cap C \neq \varnothing$ then the belief m is false, while if $P \not\subseteq C$ then you do not have all attitudes in P, hence stop reasoning prematurely. In both cases, you never get to the point of forming (and later losing) the belief m. In sum, the rule $(P \cup \{m\}, k)$ applies in the following non-standard way to any constitution C. If $K \cap C = \varnothing$ and $P \subseteq C$, then C is first transformed into $C \cup \{m\}$, which is then transformed by the rule into $C \cup \{m, k\}$, which is then transformed into $(C \cup \{k\})\backslash\{m\}$.[36] Otherwise, C is not transformed.

This model is hardly parsimonious – a drawback of modelling the full psychological process postulated by the fourth account. Reasoning on the fourth model is effectively equivalent to reasoning on the other models or the shorthand model, because the non-standard rules of the fourth model produce the same result as the non-standard rules of the second account, and hence as the indeterministic rule of the shorthand model. To be precise, this effective equivalence holds with respect to all attitudes *except* the introspective belief m. Indeed, reasoning in the other models never affects the presence of m, whereas reasoning in the fourth model can have a (final) effect on the presence of m.[37]

D Proof of Theorem 1

For any constitution $C \subseteq M$, reasoning system S, and number $n \in \{0, 1, ...\}$, let $Cn_{S,n}(C)$ denote the set of constitutions reachable from C in n steps of S-reasoning. Now fix a reasoning system S. The corresponding operator Cn_S is obviously inclusive and idempotent, for reasons already indicated. To show monotonicity, consider constitutions $C, D \subseteq M$ such that $C \subseteq D$, and fix a $D^+ \in Cn_S(D)$. Let S' be the reasoning system arising from S by replacing each rule $(P, K) \in S$ satisfying $K \cap D^+ \neq \varnothing$ with the rule $(P, K \cap D^+)$. So,

$$S' = \{(P, K) \in S : K \cap D^+ = \varnothing\} \cup \{(P, K \cap D^+) : (P, K) \in S, K \cap D^+ \neq \varnothing\}.$$

Claim 1: For every number of steps $n \in \{0, 1, 2, ...\}$, each constitution $C^+ \in Cn_{S',n}(C)$ satisfies $C^+ \subseteq D^+$ and $C^+ \in Cn_{S,n}(C)$.

We prove this claim by an induction on n. For $n = 0$, then the claim is obvious because $Cn_{S',0}(C) = Cn_{S,0} = \{C\}$ and because $C \subseteq D \subseteq D^+$. Now assume $n > 0$ and suppose the claim holds for smaller numbers than n. Fix any $C^+ \in Cn_{S',n}(C)$. Pick a

[36] Note that $(C \cup \{k\})\backslash\{m\} = C \cup \{k\}$ if $m \notin C$.

[37] One might at first think that even in the fourth model reasoning has no *final* effect on m, since m is removed after having been added. This conclusion is however hasty, because the initial constitution C might already contains m, so that 'adding' m means 'keeping' m. Here reasoning does have a final effect on m: m was initially present, but is finally absent.

$\tilde{C} \in Cn_{S',n-1}(C)$ such that $C^+ \in Cn_{S',1}(\tilde{C})$. The fact that $C^+ \in Cn_{S',1}(\tilde{C})$ and the definition of S' imply that $C^+ \subseteq \tilde{C} \cup D^+$. So, as by induction hypothesis $\tilde{C} \subseteq D^+$, we have $C^+ \subseteq D^+$.

It remains to show that $C^+ \in Cn_{S,n}(C)$. As $\tilde{C} \in Cn_{S,n-1}(C)$, it suffices to prove $C^+ \in Cn_{S,1}(\tilde{C})$. As $C^+ \in Cn_{S',1}(\tilde{C})$, we can pick a rule $(P, K') \in S'$ by which C^+ arises from $\tilde{C}$. By definition of S', there exists a $K \subseteq M$ such that $(P, K) \in S$ and *either* $[K \cap D^+ = \varnothing$ and $K' = K]$ *or* $[K \cap D^+ \neq \varnothing$ and $K' = K \cap D^+]$. The first case is impossible: it would imply that D^+ is not closed under S-reasoning, because S contains the rule (P, K) which modifies D^+ since $P \subseteq D^+$ (as $P \subseteq \tilde{C}$ and $\tilde{C} \subseteq D^+$) and since $K \cap D^+ = \varnothing$. So the second case holds. Note that

$$\tilde{C} \cap K = \tilde{C} \cap [K' \cup (K \backslash D^+)] = (\tilde{C} \cap K') \cup (\tilde{C} \cap (K \backslash D^+)).$$

In the last expression, $\tilde{C} \cap K' = \varnothing$ (as otherwise the rule (P, K') could not change $\tilde{C}$) and $\tilde{C} \cap (K \backslash D^+) = \varnothing$ (as $\tilde{C} \subseteq D^+$ by induction hypothesis). So $\tilde{C} \cap K = \varnothing$. Since $P \subseteq \tilde{C}$ and $\tilde{C} \cap K = \varnothing$, the rule (P, K) applies to $\tilde{C}$, just as the rule (P, K'). So, one can reason from $\tilde{C}$ to C^+ not just using (P, K'), but also using (P, K). In other words, C^+ belongs not just to $Cn_{S',1}(\tilde{C})$, but also to $Cn_{S,1}(\tilde{C})$. Q.e.d.

Claim 2: $Cn_{S'}(C) \subseteq Cn_S(C)$.

Let $C^+ \in Cn_{S'}(C)$. Then $C^+ \in Cn_{S',n}(C)$ for some $n \in \{0, 1, ...\}$. So, by Claim 1, $C^+ \in Cn_{S,n}(C)$. It remains to show that C^+ is stable under S-reasoning. To show this, consider any rule $(P, K) \in S$ such that $P \subseteq C^+$. We must show that $K \cap C^+ \neq \varnothing$. Form the rule $(P, K') \in S'$, where K' is K if $K \cap D^+ = \varnothing$ and is $K \cap D^+$ otherwise. Since C^+ is closed under S'-reasoning, the rule (P, K') does not change C^+, i.e., $K' \cap C^+ \neq \varnothing$. So, $K \cap C^+ \neq \varnothing$. Q.e.d.

Claim 3: Some $C^+ \in Cn_S(C)$ satisfies $C^+ \subseteq D^+$.

Pick a $C^+ \in Cn_{S'}(C)$. By Claim 2, $C^+ \in Cn_S(C)$. Further, $C^+ \in Cn_{S',n}(C)$ for some n, and so $C^+ \subseteq D^+$ by Claim 1. Q.e.d.

References

Beall, J. (2011). Multiple-conclusion LP and default classicality. *The Review of Symbolic Logic*, 4(2), 326–336. https://doi.org/10.1017/S1755020311000074

Boghossian, P. (2014). What is Reasoning? *Philos Stud*, 169, 1–18. https://doi.org/10.1007/s11098-012-9903-x

Broome, J. (2006). Reasoning with preferences? In: *Preferences and Well-Being*, S. Olsaretti ed., Cambridge University Press, 2006, pp. 183–208. https://doi.org/10.1017/CBO978051

Broome, J. (2013). *Rationality Through Reasoning*. Hoboken: Wiley. https://doi.org/10.1002/97811186

Cozic, M., Hill, B. (2015). Representation theorems and the semantics of decision-theoretic concepts, *Journal of Economic Methodology*, 22(3), 292–311. https://doi.org/10.1080/1350

Dietrich, F., List, C. (2016). Mentalism versus Behaviourism in Economics: A Philosophy-of-science Perspective. *Economics & Philosophy*, 32(2), 249–281. https://doi.org/10.1017/S0266267

Dietrich, F., List, C. (2021). The relation between degrees of belief and binary beliefs: a general impossibility theorem. In: I. Douven (ed.) *Lotteries, Knowledge, and*

Rational Belief. Essays on the Lottery Paradox. Cambridge University Press, 2021, pp. 223–54. https://doi.org/10.1017/9781108379755.012

Dietrich, F., Staras, A., Sugden, R. (2019). A Broomean model of rationality and reasoning. *Journal of Philosophy*, 116, 585–614. https://doi.org/10.5840/jphil20191161138

Dietrich, F., Staras, A., Sugden, R. (forth.). Abstract rationality: the 'logical' structure of attitudes. *Economics and Philosophy*.

Drucker, D. (2021). Reasoning beyond belief acquisition. *Noûs*. https://doi.org/10.1111/nous.12363

Felappi, G. (forth.). Propositionalism and Questions that do not have Correct Answers. *Erkenntnis*. https://doi.org/10.1007/s10670-021-00442-5

Guala, F. (2019). Preferences: neither behavioural nor mental. *Economics & Philosophy*, 35(3), 383–401. https://doi.org/10.1017/S0266267118000512

Halpern, J. Y. (2017). *Reasoning About Uncertainty*. Cambridge, Massachusetts & London: MIT Press. https://doi.org/10.7551/mitpress/10951.001.0001

Harman, Gilbert (1984). Logic and reasoning. *Synthese*, 60(1),107–127. https://doi.org/10.1007/BF004

Hausman, D. (2012). Preference, Value, Choice, and Welfare. *Cambridge University Press*. https://doi.org/10.1017/CBO9781139058537

Kolodny, N. (2005). Why be rational? *Mind*, 114, 509–563. https://doi.org/10.1093/mind/fzi509

Kolodny, N. (2007). State or process requirements? *Mind*, 116, 371–385. https://doi.org/10.1093/mind/

Kyburg, H. E., Jr. (1961). *Probability and the Logic of Rational Belief*. Middletown, CT: Wesleyan University Press. https://doi.org/10.1086/288160

Liu, F. (2011). *Reasoning About Preference Dynamics*. Dordrecht: Springer.

Montague, M. (2007). Against Propositionalism. *Noûs*, 41, 503–518. https://doi.org/10.1111/j.1468-0068.2007.00657.x

Perea, A. (2012). *Epistemic Game Theory: Reasoning and Choice*. Cambridge University Press. https://doi.org/10.1017/CBO9780511844072

Restall, G. (2005). Multiple conclusions. In Hájek, P., Valdes-Villanueva, L., & Westerstahl, D. (Eds.), Logic, Methodology and Philosophy of Science: Proceedings of the Twelth International Congress (pp. 189–205). London: King's College Publications.

Staffel, J. (2013). Can there be reasoning with degrees of belief? *Synthese*, 190(16), 3535–3551. https://doi.org/10.1007/s11229-012-0209-5

Tarski, A. (1956). *Logic, semantics and metamathematics*. Oxford University Press. https://10.1017/S0008439500024814