RETHINKING GIBBARD’S RIVERBOAT ARGUMENT

This chapter is based on a joint paper with Sylvia Wenmackers and Igor Douven published in Studia Logica (Krzyżanowska et al. 2014) and on a paper included in the book “Philosophical and Formal Approaches to Linguistic Analysis,” edited by Piotr Stalmaszczuk (Krzyżanowska 2012).

5.1 CONDITIONAL NON-CONTRADICTION

According to the principle of Conditional Non-Contradiction (CNC), conditionals with the same antecedent and contradictory consequents cannot both be true, unless the antecedent is inconsistent.¹ We follow the mainstream in regarding this principle as an adequacy constraint on any truth-conditional semantics for conditionals. After all, it may be reasonable to believe that

(88) If Matt is on holidays, he is in Venice.

or that

(89) If Matt is on holidays, he is not in Venice.

but, it would seem, we could at no one time reasonably believe both. That immediately rules out the material conditional account, which assigns to a conditional the truth conditions of its material counterpart (see section 2.2 for a discussion of this theory). Given this semantics, “If \( \varphi \), \( \psi \)” and “If \( \varphi \), not \( \psi \)” are both true whenever \( \varphi \) is false. But surely CNC leaves many other truth-conditional semantics in the running? An old but still influential argument by Gibbard (1981) purports to show that the answer is negative. In Gibbard’s view, any truth-conditional semantics that validates CNC is bound to get things wrong—from which he concludes that conditionals do not have truth conditions.

Central to Gibbard’s argument is this story:

¹ Some authors use the label CNC to refer to what we call CNC minus the proviso that the antecedent be consistent. The principle as stated here is then referred to as “Restricted Conditional Non-Contradiction.” See for instance Unterhuber (2013, Ch. 3), which also contains a thorough discussion of the status of both principles in the context of various semantics for conditionals.
Sly Pete and Mr. Stone are playing poker on a Mississippi riverboat. It is now up to Pete to call or fold. My henchman Zack sees Stone’s hand, which is quite good, and signals its content to Pete. My henchman Jack sees both hands, and sees that Pete’s hand is rather low, so that Stone’s is the winning hand. At this point, the room is cleared. A few minutes later, Zack slips me a note which says “If Pete called, he won,” and Jack slips me a note which says “If Pete called, he lost.” I know that these notes both come from my trusted henchmen, but do not know which of them sent which note. I conclude that Pete folded. (Gibbard 1981, p. 231)

Gibbard argues that if

(90) If Pete called, he won.

and

(91) If Pete called, he lost.

have a truth value at all, they must both be true. This is because Zack and Jack both appear warranted in their assertions, and their warrants do not rest on any false beliefs about relevant matters of fact. But CNC rules that (90) and (91) cannot both be true. Truth-conditional semantics for conditionals that validate CNC—as we granted they must—cannot get this right.

Or rather, Gibbard does see an escape route for the advocates of such semantics, namely, to claim that conditionals have their meanings relative to speakers’ epistemic states, so that (90) does not express the same proposition when it is uttered by Zack as when it is uttered by Jack. Supposing this to be true, we would equivocate by taking CNC to apply to (90) and (91). Stalnaker (1984, p. 109), the principal addressee of Gibbard’s critique, seconds this solution to the riverboat puzzle and compares the case of the above conditionals with an unproblematic instance of context-dependence:

There is a straightforward sense in which the sentences “I was born in Istanbul” and “I was not born in Istanbul” are contradictories: when both are interpreted relative to the same context they must have opposite truth-values. But different speakers create different contexts, and so two tokens of the respective sentences
may both be true if spoken by different speakers. Since I accept what you say, I am prepared to assert the *proposition* you expressed, but I will have to use a different sentences to do so.

Jack and Zack are in different epistemic situations, hence the context in which Jack asserts (91) is different that the context in which Zack asserts (go), so these speakers do not contradict each other at all. In Stalnaker’s semantics, the selection functions are pragmatically determined. If we have a closer look at the henchmen’s epistemic situations, we can clearly see that they will consider completely different possible worlds as the closest antecedent-worlds.

There are, at least *prima facie*, four relevant possibilities to be considered:

- \( w_{cw} \): Pete calls Pete has a winning hand
- \( w_{cl} \): Pete calls Pete has a losing hand
- \( w_{fw} \): Pete does not call Pete has a winning hand
- \( w_{fl} \): Pete does not call Pete has a losing hand

Zack knows that Pete knows his opponent’s cards. He also strongly believes in Pete’s rationality and familiarity with the rules of the game. Therefore, he does not consider the world \( w_{cl} \) possible at all, and only \( w_{cw}, w_{fw} \) and \( w_{fl} \) are compatible with what he knows. In Zack’s view, Pete will call only in those possible worlds in which he has a winning hand. In Stalnaker’s terms, from Zack’s perspective in the closest (and, in fact, the only available) “Pete calls”-world, “Pete wins” is true, and therefore the conditional “If Pete called, he won” is true.

- \( w_{cw} \): Pete calls Pete has a winning hand
- \( w_{fw} \): Pete does not call Pete has a winning hand
- \( w_{fl} \): Pete does not call Pete has a losing hand

At the same time Jack, knowing the exact distribution of cards, eliminates all those worlds in which Pete has a winning hand. Hence the only worlds he considers possible are \( w_{cl} \) and \( w_{fl} \):

- \( w_{cl} \): Pete calls Pete has a losing hand
- \( w_{fl} \): Pete does not call Pete has a losing hand

As a result the closest “Pete calls”-world selected by Jack is \( w_{cl} \) in which Pete has a losing hand. Clearly, the sets of possible worlds that constitute epistemic states of the speakers are different and
their respective selection functions take different values. “If Pete
called, he won” can indeed express two different propositions de-
pending on the context, that is, on whether it is uttered by Zack or
by Jack. But—Gibbard thinks—this is a kind of context-sensitivity
that is hard to uphold. After all, we do seem able to make sense
of (90) and (91) even if we are ignorant of the epistemic states that
the authors of these notes were in at the time of their writing.

5.2 PRAGMATIC AMBIGUITY

We already granted CNC. We also agree with Gibbard’s claim that
if one of (90) and (91) is true, then so is the other. If we still do not
accept the conclusion that conditionals must lack truth conditions,
then this is because we disagree with Gibbard’s appraisal of the
escape route he mentions. Why should it be more objectionable
to claim that a conditional can express different propositions in
the mouths of different speakers than to claim that, for instance,

(92) I am the owner of this house.

or

(93) Everyone has got a drink.

express different propositions in the mouths of different speak-
ers? (92) may be true when said by Sally and false when said by
Paul, yet the two speakers clearly do not contradict each other.
The truth value of (93) depends on the intended scope of the the
quantifier and, again, the reference of the indexical “here.” It can
be true when uttered by Clara at her house warming party and
false when uttered by Robert at a rock concert. According to Gib-
bard, however, the cases are not parallel because we can rely on
fairly straightforward rules for detecting the relevant contextual
information necessary to interpret sentences like (92) or (93): “‘I’
refers to the speaker, ‘this’ to whatever the speaker points at, ‘here’
refers to the place of utterance.” Similarly, the scope ambiguous
quantifier “everyone” can be easily interpreted as ranging over a
group of people salient in the context of utterance.

---

2 Our agreement is based on pre-theoretical intuition rather than on Gibbard’s ar-
argument for this claim. True, neither Jack’s nor Zack’s assertion rests on a mistaken
belief, but Jack, in contrast to Zack, still lacks knowledge of a seemingly highly
relevant matter of fact, to wit, what cards Pete is holding. While Zack may not
have any mistaken beliefs about these cards, his ignorance might still lead him to
assert something false—even warrantedly so, given the quality of the evidence
that is in his possession. See, in the same vein, Lycan (2001, p. 169).
Gibbard (1981, p. 232) argues that the difference in contexts in which (90) and (91) occur has “a strange feature,” because:

Ordinarily when context resolves a pragmatic ambiguity, the features of the context that resolve it are common knowledge between speaker and audience. If the chairman of a meeting announces “Everyone has voted ‘yes’ on that motion,” what the audience knows about the context allows it to judge the scope of ‘everyone’.

In the sly Pete story, by contrast:

whatever contextual differences between the utterances there may be, they are unknown to the audience. I, the audience, know exactly the same thing about the two contexts: that the sentence is the content of a note handed me by one of my henchmen. Whatever differences in the context make them invoke different s-functions is completely hidden from me, the intended audience.

Gibbard concludes that there do not seem to exist similarly straightforward rules for interpreting context-sensitive conditionals. Stalnaker, who not only acknowledges but also endorses the context-sensitivity of conditional sentences, admits that asserting a conditional, especially in contexts resembling the one described by Gibbard, seems to be a violation of a conversational maxim that teaches us to assume that our interlocutors have all the information necessary to interpret our messages.

In general, in any context in which If A then B and In A then not-B are (epistemically) possible answers, the questioner will have the contextual information necessary to interpret the answer only when he knows what it is. This seems to violate the following conversational maxim: speakers ought, in general, to assume that their addressees have whatever information is necessary to determine what they are saying. It seems that the assumption that open conditionals express propositions leads to the conclusion that in a very general kind of context, the assertion of indicative conditionals will always conflict with a very plausible principle of conversation. (Stalnaker 1984, pp. 110-111)
Stalnaker’s advice is to relax the maxim so it allows for exceptions. But this is not really a solution to the problem that Gibbard is bothered with, namely, that if conditionals do express propositions, these propositions are extremely subjective, that is, they depend on the epistemic states of the speakers who utter them, and we rarely have direct access to what other speakers know or believe. And this is exactly the kind of context-sensitivity Gibbard finds unacceptable.

We dispute Gibbard’s assessment of the situation and believe that, pace Gibbard, interpreting context-sensitive conditionals is not so different from interpreting sentences containing indexicals. This is not to deny that we can make sense of (90) and (91) even if we have no idea which note is from Zack and which from Jack, nor know much about Zack’s and Jack’s belief states. But in the same way we can make sense of (92) or of (93) if these are written on notes of unknown origin, and we do not know the referents of “I” or “this house,” in the first case, or the scope of the quantifier “everyone,” in the second case. What is true is that in such cases we cannot fully interpret (92) nor (93), but we still get something from them. We get from (92) that whoever wrote the message claims to be the owner of whichever house “this house” is meant to refer to. In some contexts, we might even be able to say with some confidence who the author is and to which house he or she is referring. A message containing a scope-ambiguous quantifier “everyone” would be particularly hard to interpret. If I do not know anything about the context in which (93) was uttered, I will have no clue what is the scope of the quantifier occurring there. However, from (93), we can get that each member of a group of people present in the place where the message was written has got a drink. In our view, much the same is true for (90) and (91). Having been given the notes with (90) and (91) on them, but not knowing who gave us which note and what the persons’ epistemic states were, we can still get something out of those messages. To be sure, one gets more out of the messages if one does know who handed one which note, and what the relevant beliefs of the authors were at the time of writing. But that is as with the notes with (92) or (93) on them: one gets more out of those if one does know the intended referents of “I,” “this house,” and “everyone” than if one does not.

In the following section, we will argue that the truth-conditional semantics for conditionals introduced in chapter 3 renders both (90) and (91) true, given the specifics of the above story. Moreover, the position allows us to keep CNC on board while avoiding the
untoward consequences that Gibbard envisions for any semantics that factors a speaker’s epistemic state into the truth conditions of conditionals. Additionally, in section 5.5 we report some experimental evidence favouring our diagnosis of the Gibbard story.

5.3 A NEW SOLUTION

In chapter 3, we proposed a new semantic theory of conditionals based upon an idea that many, if not all, conditional sentences convey the existence of an inferential connection between the content of their antecedent and the content of their consequents. Importantly, we acknowledged the variety of possible consequence relations linking antecedents and consequents, following the typology of inferential conditionals introduced by Douven and Verbrugge (2010). We defined the truth conditions in the following way:

**Definition 3:** A speaker S’s utterance of “If φ, ψ” is true if and only if:

(i) ψ is a consequence of φ in conjunction with S’s background knowledge,

(ii) ψ is not a consequence of S’s background knowledge alone but not of φ on its own, and

(iii) φ is deductively consistent with S’s background knowledge or ψ is a consequence of φ alone,

where the consequence relation can be deductive, abductive, inductive, or mixed.

To see what follows from Definition 3 for the Gibbard story, first consider a variant of that story. Sly Pete and Mr. Stone are playing poker again, with Zack and Jack again doing their business. Once more, it is up to Pete to call or fold, who also in this story has been signalled the content of Mr. Stone’s hand by Zack. In contrast to the original story, however, Pete now has the upper hand. Zack does not know this, but Jack, who has again seen both players’ hands, does. After the room has been cleared, both Zack and Jack sincerely assert:

(90) If Pete called, he won.

Although by doing so, they both assert something true, what makes Zack’s assertion true is not the same as what makes Jack’s
assertion true. Relative to Zack’s background knowledge, the best explanation for Pete’s calling—supposing he called—is that he held cards that were better than Mr. Stone’s, which Pete knew thanks to Zack’s fraudulent act. From this in conjunction with the antecedent of (90), it follows deductively that Pete won, so that, from Zack’s perspective, the consequent of (90) is a mixed consequence of the antecedent of this conditional. That is what makes Zack’s assertion of (90) true. On the other hand, relative to Jack’s background knowledge, that Pete won follows logically from his calling together with the information about both players’ hands and the rules of poker. That is what makes Jack’s assertion of (90) true.

The proposed diagnosis of what is going on in Gibbard’s original story will now be unsurprising. From Jack’s perspective, “Pete lost” obviously follows deductively from “Pete called,” which is what makes his assertion of (91) true. This does not mean that Zack’s utterance of (90) must be false. To the contrary, as in the above story, from Zack’s perspective Pete’s calling is best explained by his having a winning hand, from which, in conjunction with the antecedent, it follows deductively that Pete won—and this is what makes Zack’s assertion of (90) true. That is to say, according to Definition 3, both Jack and Zack assert something true.

A number of authors (e.g., Bennett 2003, p. 85 or Edgington 1995, p. 294-295) have presented “symmetric” versions of Gibbard’s story which involve pairs of conditionals that are both of the same inferential type, for instance:

Suppose there are two vaccines against a certain disease, A and B. Neither is completely effective against the disease. Everyone who has A and gets the disease, gets a side effect S. Everyone who has B and gets the disease, doesn’t get S. Having both vaccines is, however, completely effective against the disease (though not many people have both). These scientific facts are known. X knows that Jones has had A, and says “If Jones gets the disease, he’ll get S.” Y knows that Jones has had B, and says “If he gets the disease, he won’t get S.” (Edgington 1995, p. 294-295)

But given that in those versions the personae asserting the conditionals still have different background knowledge, there is no problem acknowledging that the conditionals are both true. More precisely, there is nothing wrong with “If φ then ψ” being true
because $\psi$ follows from $\varphi$ in conjunction with a set of background premises $\{\alpha_1, \ldots, \alpha_n\}$, while, at the same time, “If $\varphi$ then not $\psi$” is true because “not $\psi$” follows from $\varphi$ in conjunction with a different set of background premises, $\{\beta_1, \ldots, \beta_n\}$.

On the other hand, it might be said that accounts like the one presented in this thesis, that relativise the truth of a conditional to the speaker’s background knowledge, fail to address the intuition that Zack and Jack disagree with each other. In our view, however, their disagreement is only apparent. If Jack and Zack were to talk to each other, they would quickly realize that they have different information, which might lead them to share what they know. Probably, that would make Zack change his mind. But this would not really amount to retracting his earlier assertion. Rather, we would expect Zack to say something like the following once Jack had revealed his information: “What I meant to say is that if Pete called, that would have indicated that he had the upper hand. But now that Jack told me that in fact Mr. Stone had the upper hand, I’m sure Pete did not call.” After they have exchanged their information, Pete’s calling becomes inconsistent with Zack’s and Jack’s shared background knowledge. As observed by, for instance, Edgington (1995, p. 295):

The Gibbard phenomenon arises if and only if there are currently ascertainable facts that rule out [the antecedent].

If it turns out that, in the original version of the story, Pete actually called, it would mean that Zack was actually mistaken about something, for instance, he wrongly believed that Pete played to win or that he correctly interpreted Zack’s signals. If that were the case, the consequent of Zack’s conditional would not be a consequence—be it deductive, inductive, abductive or mixed—of Zack’s background knowledge, and, consequently, (90) would be false. What Zack could truthfully assert in such a situation, however, given his background knowledge, could be, for instance, another abductive inferential conditional, “If Pete called, he does not want to win” or “If Pete called, he is irrational.”

In summarizing Gibbard’s argument in Section 5.1, we said that, by taking the escape route suggested by Gibbard, his story would no longer present a violation of CNC. We took that escape route, but it could be objected that, strictly speaking, we still have a violation of CNC. For, as it stands, CNC assumes that one can determine merely on the basis of form whether a conditional in the mouth of one speaker is inconsistent with a conditional in
the mouth of another speaker. And the above proposal has not done anything to alter the form of either (90) or (91). As will be clear, we take this to overlook the crucial role that the background knowledge of a speaker plays in determining the interpretation of a conditional. To make this role explicit, it suffices to reformulate CNC as the principle that utterances of instances of “If $\varphi$, $\psi$” and “If $\varphi$, not $\psi$” cannot both be true in conjunction with the same body of knowledge. Put this way, the conditionals in Gibbard’s story no longer constitute a violation of CNC.

To come to what is arguably the main challenge of Gibbard’s argument, we note that while our view makes the interpretation of conditionals relative to speakers’ background knowledge, it does not lead to a relativism that must leave one completely clueless as to the interpretation of any given conditional. Not knowing who wrote which note, we still understand that Pete’s losing is a consequence of his calling relative to one of the henchmen’s background knowledge, and Pete’s winning is a consequence of his calling relative to the other henchman’s background knowledge. More precisely, what the message with (90) on it conveys is that, from the perspective of its author, “Pete won” follows—by virtue of deduction, induction, abduction or a combination of these—from “Pete called.” Likewise, what we learn from the message with (91) on it is that its author has grounds to infer—again, in one way or the other—“Pete lost” from the assumption that Pete called. Naturally, we understand more if we know who wrote which note, and if we know something about his background knowledge. For instance, knowing that (90) comes from Zack, and that Zack signalled Mr. Stone’s hand to Pete but did not see Pete’s hand, it should not take too much effort to figure out that Pete’s winning is not just a consequence of his calling, but more specifically a mixed consequence, involving an abductive inferential step. Still, Gibbard or anyone else defending his line would be hard pressed to argue how any of this is essentially different from reading a note with, for instance, (92) on it in a context in which one is ignorant of who wrote the note as compared to reading that note in a context in which one does know who wrote it. They would be equally hard pressed to argue that, on our account, one gets less from reading (90) or (91) than one does from a pre-theoretic viewpoint.

The main general point to be emphasized in this connection is that interpretation is not a matter of all or nothing. It is not as though we always either do or do not get what a speaker says; there are many shades of partial understanding in-between these
extremes. By reading his henchmen’s notes, the first-person narrator in Gibbard’s story gets a lot, just not everything.

5.4 DELIBERATIONALLY USELESS CONDITIONALS

Gibbard’s poker game scenario does not only pose a challenge to propositional theories of conditionals that validate the Principle of Conditional Non-Contradiction (CNC), but it can also serve as a case against indicative conditionals-based decision making. Incidentally, contexts of deliberation provide some extra motivation for endorsing CNC: a piece of advice of the form “If you take a cab, you will not be late for the concert” would be useless if “If you take a cab, you will be late for the concert” can be truthfully asserted in the same context, given the same background knowledge. The semantics proposed in this thesis presents a solution to the problem raised by Gibbard (see section 5.3), and, additionally, sheds new light on the phenomenon of so-called deliberationally useless conditionals.

It is a common assumption that conditionals of deliberation, that is conditionals that a rational agent considers in the process of decision making, are future directed counterfactuals of the form “If I were to do \(\phi\), then \(\psi\) would happen” (see e.g. Gibbard and Harper 1978 and Joyce 1999). For instance, someone who recently got into financial troubles may think:

(94)  
   a. If I were not to ask my sister for a loan, I would have to go to a bank.
   
   b. If I were to borrow money from a bank, I would have to pay the interest rate.

Clearly, realising the dependencies between the states of affairs expressed by the above conditionals’ antecedents and their consequents can help the agent to make a right decision. But could he not phrase them in a more straightforward way as “If I do \(\phi\), then \(\psi\) will happen”? At least \textit{prima facie}, the following future directed indicatives:

(95)  
   a. If I don’t ask my sister for a loan, I will have to go to a bank.
   
   b. If I borrow money from a bank, I will have to pay the interest rate.
can act as conditionals of deliberation, too. Sentences in both (94) and 95 seem to be equally closely related to yet another type of conditionals that can guide agents in their decision making:

(96) a. If I don’t want to pay the interest rate, I shouldn’t borrow money from a bank.

b. If I don’t want to go to a bank, I should ask my sister for a loan.

Sentences of the form “If I want \( \psi \), I should do \( \varphi \)” can be classified as “goal-directed” conditionals of deliberation.

Despite the close relationship between grammatically subjunctive conditionals in (94) and their indicative counterparts in (95), it seems that there are reasons not to follow indicative conditionals in contexts of decision making. Most notably, some indicatives can constitute bad or even conflicting advice, and hence they can mislead a deliberating agent. Yet DeRose (2010) maintains that conditionals of deliberation are in fact indicatives, albeit some of them are unassertable in contexts of deliberation. However, although I agree with DeRose’s main claim, I am going to show that he failed to recognise the reason why certain indicative conditionals should not be used in contexts of decision making. Furthermore, I will argue that the inferential semantics of conditionals proposed in chapter 3 offers an alternative explanation of the phenomenon of bad or conflicting advice.

### 5.4.1 The problem of bad advice

DeRose (2010) observed that Gibbard’s poker game scenario discussed in the previous sections is a good example of a context in which indicative conditionals may pose a problem to a deliberating agent.³ To recapitulate the scenario, Pete is playing poker against Mr. Stone. Zack signals the content of Stone’s hand to Pete, while Jack peeks at both player’s hands and sees that Stone has the winning hand. The only aspect of Gibbard’s original story that needs to be changed for it to become a context of deliberation are the tenses of the sentences asserted by Zack and Jack. These conditionals, originally in the past tense, have to be rephrased as future directed conditionals. Accordingly, before Pete makes his decision, Zack sends Gibbard a message saying:

(97) If Pete calls, he will win.
As has been argued earlier in this chapter, this is a warranted assertion, given that Zack knows that Pete is cheating, and hence will not call unless he has the winning hand. Yet (97) would constitute a very bad guideline for Pete in the process of deciding whether to play or to quit. Similarly, even though driving a Maserati is very good evidence for being rich, and it is perfectly reasonable to believe that if Charlie drives a Maserati, he is rich,

\( (98) \) If Charlie drives a Maserati, he will be able to afford a stay in a 5 star hotel.

taken by Charlie as a piece of advice may lead to disastrous consequences. Since if Charlie, who does not own a Maserati, nor can afford a 5 star hotel, overhears (98) and interprets it as an instruction on how to get rich, it could lead him to take an enormous loan for a luxurious car. That would not only be an irrational decision to make, but also extremely counter-productive, given Charlie’s goal of becoming rich. Analogously, Pete interpreting (97) as a guideline can end up calling despite his unfavourable position in the game.

By contrast, Jack’s assertion:

\( (99) \) If Pete plays, he will lose.

is based on his knowledge of exactly what cards both players have. On that account, it would not only be a piece of advice conflicting with the one from Zack, but, at the same time, it would be an extremely valuable one. This observation alone is not helpful though, unless we know how to accurately tell those conditionals that will not mislead us from those that will or at least may mislead us. Without such a criterion, one seems to be better off avoiding following pieces advice phrased as indicative conditionals.

Fortunately, in the cases like (97) and (98), as DeRose points out, a deliberating agent would not consider the indicative conditional as an instruction if he knew what are the speakers’ grounds for asserting it. Likewise, should the speakers of (97) and (98) realise that they are involved in the contexts of Pete’s and Charlie’s decision making, respectively, they would never assert (97) and (98) (or at least not without any explanations). And the reason for this, as DeRose argues, is that conditionals like (97) and (98) are simply deliberationally useless and hence unassertable in the context of deliberation, whereas (99) is an example of a deliberationally useful conditional.
As noted, acknowledging the distinction between deliberationally useless and deliberationally useful conditionals would not be of much help, if we had no systematic way to recognise a conditional’s membership in one of the classes. DeRose’s own proposal is that future directed conditionals “are deliberationally useless when they are based on backtracking grounds” (pp. 28). Indeed, in the poker game scenario, Zack supposes that Pete calls, and then he reasons back in time to realise that Pete knows the opponent’s card and hence will play only when his cards are stronger. But is it really the backtracking that we should avoid when involved in the process of decision making?

Although DeRose’s analysis of Zack’s grounds for asserting (97) is basically correct, he still fails to recognise the reason why (97) or (98) are deliberationally useless. Although they both hinge upon backtracking reasoning, it is not backtracking itself that renders those conditionals unassertable in contexts of deliberation. In fact, I am going to argue that a whole class of deliberationally useful indicative conditionals notably depends on reasoning back in time.

5.4.2 Backtracking in the context of deliberation

As mentioned above, goal-directed conditionals of the form “If I want $\psi$, I should do $\varphi$” can be thought of as just another way of expressing dependencies between certain actions and states of affairs that are crucial for rational decision making. Moreover, sentences like those in (96) or:

(100) a. If Bob wants to have a well-paid job, he should study law.

b. If you don’t want to be late for the meeting, you should leave in 10 minutes.

sound in contexts of deliberation perfectly natural. Assuming that their assertions are well justified, they usually constitute rather good pieces of advice. And yet they involve some form of backtracking. Though they are not backtracking conditionals per se, and the reasoning pattern here is slightly different than in the cases discussed by DeRose, their dependence on reasoning back in time is significant. The antecedents of the above conditionals are propositional attitude reports, and as such they do not seem to be preceded by whatever is expressed by the consequents. But to figure out what should be done in order to fulfill the wishes
or plans reported by the antecedents, one has to start with the objects of those propositional attitudes.

For instance, the speaker of (100a) considers the possibility that Bob finds a well-paid job and reasons back in time to find out what may make this happen, that is, which study programme would most likely lead to a career providing a good pay. He arrives at the conclusion that lawyers usually earn good money, and hence that Bob should study law. In other words, one can follow the procedure below:

1. Assume the antecedent of the conditional, \( \varphi \) (e.g. “Bob wants to have a well-paid job”);
2. “Extract” the embedded sentence \( \varphi^* \) (e.g. “Bob has a well-paid job”) from the attitude report “an agent \( a \) wants \( \varphi^* \)” (“Bob wants that Bob has a well-paid job”);
3. Assume \( \varphi^* \);
4. Reason back in time to see what could lead \( \varphi^* \) to be the case. Given the relevant background knowledge, choose what is the best way to achieve \( \varphi^* \), and call it \( \psi^* \) (e.g. “Bob studies law”).
5. Rephrase \( \psi^* \) as an instruction \( Q \) for \( a \) by adding a modal verb “should” to \( \psi^* \), as in “an agent \( a \) should \( \psi^* \)” (e.g. “Bob should study law”).
6. Assert “If \( \varphi \), then \( \psi \)” (100a).

Although the backtracking reasoning involved in goal-directed conditionals is different than what DeRose is concerned with, it is still a backtracking reasoning. The example serves as an illustration that it is not reasoning back in time itself that renders conditionals like (97) deliberationally useless. Equating backtracking conditionals with deliberational uselessness would exclude from the context of decision making a vast class of conditionals that seem to be rather useful. Furthermore, saying that deliberationally useless conditionals depend on backtracking reasoning, even if true, does not explain why they are useless. Why is it the case then that some indicative conditionals that depend on backtracking are unassertable in the context of decision making? The semantics proposed in this thesis provides a straightforward answer.

\footnote{Note that studying law should be the best way to a well-paid job for Bob, given his personal qualities or the situation he finds himself in. It does not need to be the best choice for an arbitrary person with the same goal. Moreover, if there is more than one way to achieve the goal, and they are all equally good in given circumstances, the speaker is not warranted in asserting the conditional.}
5.4.3 The source of unassertability

In section 5.3 we have argued that what distinguishes Zack’s deliberationally useless conditionals from the one asserted by Jack is the type of inference those conditionals reflect. In the case of (99), on the one hand, assumption of the antecedent “Pete calls” together with background information on the distribution of cards and on the rules of the game leads to the conclusion expressed by the consequent: “Pete will lose.” It does so, because Pete’s losing follows deductively from the supposition that he called together with the background knowledge. Zack’s assertion, on the other hand, is based on backtracking reasoning (as DeRose quite correctly observed), that is reasoning on what precedes the state of affairs expressed by the antecedent. This backtracking reasoning is, as a matter of fact, an instance of abduction. Supposing that Pete calls leads Zack to the best explanation of Pete’s decision to call. In other words, what Zack reasons about is what could make Pete decide to call. The consequent of the conditional asserted by Zack is an abductive consequence of its antecedent in conjunction with Zack’s background knowledge.

How does it relate to the issue of deliberational uselessness? Simple as that, abductive reasoning, or reasoning to the best explanation, does not provide us with any information that we could use to make a rational decision. If I am trying to decide whether I should do A, what I am interested in are the consequences of the action A, not its causes or explanations. Someone explaining to me what my decision to do A would be an indication or a symptom of would not help me to make up my mind. To illustrate this point, let us consider another example. Suppose that I received an invitation to a party from people that I do not like. To the best of my knowledge, there will be no one at that party that I would really like to hang out with. On the other hand, I have no other plans for the evening, and I am longing for some company. Should I go to that party? I may consider a conditional: “If I go to that party, I am desperately lonely.” Obviously, I do not want to be desperately lonely, but does it mean that I should reject the invitation? No, because my loneliness would not be the result of my participation in the party, but rather a mere explanation of my own, somewhat surprising decision. This abductive inferential conditional does not inform me whatsoever about the benefits of either accepting or rejecting the invitation, it only provides me with an explanation of a hypothetical action that I still have no reason not
to undertake. By contrast, an inductive inferential conditional, for instance “If I go to that party, I will meet some new people” could be exactly what I need in order to make a rational decision regarding this one particular evening of my life.

Taking into account the typology of inferential conditionals in the analysis of Gibbard’s poker game scenario not only allows to identify those conditionals that are deliberationally useless, but also provides a reason and an explanation for their infamous role in the process of decision making.

5.5 EXPERIMENT: GIBBARD STORIES

Our diagnosis of what is going on in Gibbard’s argument is so far mainly supported by our own pre-theoretic responses to (90) and (91) in the context provided by Gibbard’s story. While in our opinion intuition-based approaches still have their value in philosophy, we have considerable sympathy for the recent trend to try and provide more robust, less subjective support for one’s philosophical claims and analyses. This section presents an experiment that is meant to offer precisely that kind of additional support for our analysis of Gibbard’s argument.

The experiment makes use of previous experimental work on so-called evidential inferential markers in relation to conditionals discussed in chapter 4. In that work, we found that “probably,” when occurring in the consequent of a conditional, is a good marker of conditionals that are uncertain in Douven and Verbrugge’s 2010 sense. Specifically, inserting “probably” in the consequent of a conditional has a tendency to raise that conditional’s perceived assertability if the conditional is either an II or an AI conditional as well as a tendency to lower the conditional’s perceived assertability if the conditional is a DI conditional.

This finding suggests a straightforward empirical test of our diagnosis of Gibbard’s riverboat argument. After all, according to that diagnosis, “If Pete called, he lost” is a certain conditional relative to the background information available to Jack, and so, in view of the above finding, the insertion of “probably” in the consequent should make the conditional less fitting into a context in which that information is assumed. Conversely, according to the same diagnosis, “If Pete called, he won” is an uncertain con-

5 Of course, I do not claim that abductive conditionals cannot play any role whatsoever in the process in decision making. They can, but not as conditionals of deliberation.
ditional relative to the background information available to Zack, and so the insertion of “probably” in the consequent should make the conditional more fitting into a context in which that information is assumed.

5.5.1 Method

Participants
Two hundred and seventy-eight persons participated in the experiment. They were recruited via the CrowdFlower interface (http://www.crowdflower.com), which directed them to the Qualtrics platform (http://www.qualtrics.com) on which the experiment was run. The participants were paid a small amount of money in return for their cooperation. All participants were from Australia, Canada, New Zealand, the United Kingdom, and the United States. We excluded from the analysis 81 participants who failed two comprehension questions, 11 further participants who had indicated that they were non-native speakers of English, and 3 still further participants who had spent more than 30 minutes on the survey. This left us with 183 participants. The mean age of these participants was 35 (±13); 115 participants (63%) were women; 79% indicated college or higher as their education level, 20% indicated high school, and 1% indicated a lower education level.

Design
Every participant was given five different stories and was asked which of two conditionals fitted better in the context of the story. There were two versions of each story, a “certain” and an “uncertain” one. Each participant received only one version of a story. All the materials were presented on screen, with each story together with the associated question appearing on a different screen. The order of the story–question pairs was randomized and for each separate story it was randomly determined whether the agent received the certain or the uncertain version of that story.

Materials and Procedure
All materials were in English, the participants’ native language. One group of participants (N = 83) was offered the part of Gibbard’s story containing the information available to Jack, who sees both players’ hands (which creates a “certain” context), and another group (N = 100) was offered the part of Gibbard’s story
containing the information available to Zack, who can only see Mr. Stone’s hand and signals its contents to Pete (which creates an “uncertain” context). Participants in the first group were then asked which of “If Pete called, he lost” and “If Pete called, he probably lost” they thought fitted best in the context. Participants in the second group were asked which of “If Pete called, he won” and “If Pete called, he probably won” they thought fitted best into the context. Both groups were given the options to answer that the first conditional fits best, that the second fits best, that both fit equally well, and that neither fits well. Figure 6 shows a story and a question presented on the screen to the first group of participants. The corresponding story plus question presented to the second group is shown in Figure 7.

Figure 6: Gibbard’s story, certain version, as presented to the participants.

Because our own first impression of Gibbard’s riverboat story was that it was somewhat difficult to comprehend, at least for people who are unfamiliar with the game of poker, we also included in our experiment DeRose’s 2010 version of Gibbard’s story, which is about a simpler game than poker but otherwise
Pete and Mr. Stone were playing poker yesterday. At some point, it was up to Pete to call (i.e., to continue playing, with the risk of losing if his cards are worse than Mr. Stone’s) or to fold (i.e., to quit playing). Pete is a good poker player who does not like to lose. Moreover, he was cheating: his friend, Jack, was looking at Mr. Stone’s hand and, using special signs, informed Pete about what cards Mr. Stone had.

Consider these sentences:

1. If Pete called, he won.
2. If Pete called, he probably won.

Of these sentences

- the first fits best into the context.
- the second fits best into the context.
- both fit equally well.
- neither fits well.

Figure 7: Gibbard’s story, uncertain version, as presented to the participants.

identical to Gibbard’s story.6 As we did in the case of Gibbard’s story, we split the information presented in DeRose’s story into a part containing the information available to one player’s helper—we called the player “Jim” in the survey question—who sees only what is in the hand of Jim’s opponent—called “Steve” in the survey question—and signals its content to Jim, and a part containing the information available to Steve’s helper who sees both players’ hands. Again we asked one group (N = 87) to assess the fittingness in the former context of “If Jim plays, he will win” with and without “probably” in its consequent, and another group (N = 96) to do the same for “If Jim plays, he will lose” in the latter context. (See the appendix C for the two versions of DeRose’s story that we used.)

As an additional test of the correctness of our analysis of Gibbard’s story and our hypothesis concerning what it shows about

---

6 Actually, there is a difference in the tense of the conditionals used in the two stories (see examples (97) and (99) discussed in the previous section). As far as we can see, this difference is immaterial. But of course this claim might be refuted by future empirical research.
conditionals, we invented a number of stories that, supposing our analysis to be correct, closely parallel Gibbard’s story in all relevant respects and that should therefore yield similar results with regard to the effect of inserting “probably” in the consequent of the conditionals that we supposed to parallel the conditionals in Gibbard’s story. Specifically, we came up with three stories, one about a certain Mr. Smith, one about a poisoned apple, and one about a linguistics exam, the information in each of which could be filtered in two different ways, resulting in a story that, according to our judgement, should make a conditional pertaining to the story appear uncertain, and a story that, also according to our judgement, should make the same conditional appear certain. (See again the appendix for a description of these stories.) For the certain version of the Mr. Smith story, $N = 86$; for the uncertain version, $N = 97$; for the certain version of the story about the apple, $N = 87$; for the uncertain version, $N = 96$; and for the certain version of the linguistics exam story, $N = 95$, for the uncertain version, $N = 88$.

### 5.5.2 Results

A chi-square test was performed to test the hypothesis that one of the conditionals in the Gibbard story is a certain inferential conditional and the other an uncertain inferential conditional, given that “probably” is a good linguistic marker for distinguishing between the two types of conditionals. A moderately strong association was found between type of context (certain / uncertain) and assessment of effect of insertion of “probably,” $\chi^2(3, N = 183) = 24.9$, $p < .0001$; Cramér’s $V = .369$, $p < .0001$. Inspection of the cell frequencies showed that 55 (66%) out of the 83 participants who had received the certain version of the story judged the corresponding conditional without “probably” to fit best into the context, whereas only 22 (27%) of the same 83 participants judged the conditional with “probably” to fit better; 5 of the remaining 6 participants thought both conditionals fitted equally well; 1 participant thought neither fitted well. By contrast, 51 (51%) out of 100 participants who received the uncertain version of the story judged the conditional with “probably” to fit best into the context, and only 30 (30%) of the same group of participants judged the conditional without “probably” to fit best into the context; 12 participants thought that both conditionals fitted well; 7 thought neither fitted well.
We asked participants about their familiarity with the game of poker, and an analysis of the data of participants who indicated that they knew the game of poker well (38 of the 183 participants) revealed a stronger association than the one we found for the group as a whole. In particular, Cramér’s V = .449 for this group. However, the effect size is only marginally significant (p = .053).

We obtained similar results for DeRose’s story, although here a chi-square test revealed a strong association between type of context (certain / uncertain) and effect of “probably,” \( \chi^2(3, N = 183) = 64.4, p < .0001 \); Cramér’s V = .59, p < .0001. More specifically, in the certain context, 84 (88%) out of 96 participants judged the conditional without “probably” to fit best; only 8 participants (8%) judged the conditional with “probably” to fit best; 3 thought both fitted well; 1 thought neither fitted well. On the other hand, in the uncertain context, 26 (30%) out of 87 participants judged the conditional without “probably” to fit best; 41 participants (47%) judged the conditional with “probably” to fit best; 7 thought both fitted well; 13 thought neither fitted well.

As for the story about Mr. Smith, a chi-square test again found a strong association between type of context (certain / uncertain) and effect of “probably,” \( \chi^2(3, N = 183) = 60.2, p < .0001 \); Cramér’s V = .57, p < .0001. Cell frequencies for the certain version were: 64 (74%) out of 86 participants thought that the conditional without marker fitted best; 8 (9%) thought the conditional with marker fitted best, 7 thought both fitted well; 7 thought neither fitted well. Cell frequencies for the uncertain version were: 20 (21%) out of 97 thought the conditional without marker fitted best; 56 (58%) thought the conditional with marker fitted best; 9 thought both fitted equally well; 12 thought neither fitted well.

As for the story about the apple, a chi-square test found a very strong association between type of context (certain / uncertain) and effect of “probably,” \( \chi^2(3, N = 183) = 107.9, p < .0001 \); Cramér’s V = .77, p < .0001. Cell frequencies for the certain version were: 65 (75%) out of 87 participants thought that the conditional without marker fitted best; 11 (13%) thought conditional with marker fitted best; 9 thought both fitted well; 2 thought neither fitted well. Cell frequencies for the uncertain version were: 5 (5%) out of 96 participants thought the conditional without marker fitted best; 75 (78%) thought the conditional with marker fitted best; 4 thought both fitted well; 12 thought neither fitted well.

As for the linguistics test story, a chi-square test found a strong association between type of context (certain / uncertain) and ef-
fect of “probably,” $\chi^2(3, N = 183) = 54$, $p < .0001$; Cramér’s $V = .543$, $p < .0001$. Cell frequencies for the certain version were: 50 (52%) out of 95 participants thought the conditional without marker fitted best; 30 (32%) thought conditional with marker fitted best; 9 thought both fitted equally well; and 6 thought neither fitted well. Cell frequencies for the uncertain version were: 6 (7%) of 88 thought the conditional without marker fitted best; 54 (61%) thought the conditional with marker fitted best; 4 thought both fitted equally well; and 24 thought neither fitted well.

5.5.3 Discussion

Except possibly for the apple story, the scenarios presented to the participants may have triggered some associations that we were not quite able to control. For instance, in the linguistics test story, 32% of the participants decided that the conditional with “probably” fits best in the certain context while only 53% of them thought the conditional without the marker fits best. Even though the effect of “probably” here is statistically significant, it is still curious that so many people chose the conditional with “probably.” The certainty of this scenario is based on rules that are of human origin. Participants may have their own experiences with universities, assignments and tests, and some of them might have experienced that rules of the sort at issue in the story are sometimes violated. It is well known that participants do not always fully believe what is given in vignettes and may, in view of their background beliefs, remain somewhat uncertain about what they are asked to suppose as being true (see Evans and Over (2004, Ch. 6) and Over et al. (2013)). Hence, they may not be willing to conclude $p$ on a basis of (roughly) “there is a rule that $p$,” and they may still prefer “probably $p$” in this kind of context.

The difference in the strength of association between Gibbard’s story and DeRose’s version of that story may be entirely due to the fact that the former involves the game of poker, the rules of which tend to strike outsiders as being relatively complex. Even if no real knowledge of those rules is presupposed by the story, that itself is not immediately evident unless one is at least somewhat familiar with the game. By contrast, DeRose’s version poses no particular comprehension problems. For this reason, we take the responses to DeRose’s version to be more relevant to our hypothesis, and those responses revealed a strong association between type of context and effect of inserting “probably” in the relevant
conditional. Specifically, inserting “probably” had a tendency to make the conditional appear better fitting in the uncertain context, but less fitting in the certain context. In view of the results of discussed in chapter 4, this finding is exactly what one would expect if the uncertain context invites an interpretation of the conditional as being uncertain and the certain context invites an interpretation of the conditional as being certain.

That, guided by our view on Gibbard’s story, we were able to design a number of scenarios that were meant to parallel Gibbard’s and DeRose’s and that turned out to elicit the same patterns in the responses, is further evidence for the correctness of our diagnosis.

5.6 conclusion

Gibbard’s argument challenges truth-conditional accounts of conditionals. On the one hand, CNC appears an utterly plausible principle. On the other hand, the riverboat story features two conditionals, one of the form “If \( \varphi, \psi \),” the other of the form “If \( \varphi, \neg \psi \),” which in the context of the story strike one as being both true. According to Gibbard, it follows from his story that conditionals do not express propositions, and thus are not true or false; at least this follows, Gibbard thinks, unless one is willing to buy into a semantics that makes the communicative role of conditionals hard to understand. We have argued for a different conclusion by stating a semantics that relativises the interpretation of conditionals to speakers’ background knowledge but not in a way that must leave us clueless about the meaning of a conditional whenever we are ignorant of the speaker’s background knowledge. As a consequence of our proposal, conditionals cannot be judged to be inconsistent with each other purely on the basis of their form. But this admission requires no more than a minor amendment in the formulation of CNC. Furthermore, the proposed analysis sheds new light on the role indicative conditionals can play in the context of decision making. Finally, it was seen that, in light of recent experimental results on linguistic markers, our analysis of Gibbard’s story has clear empirical content, part of which was investigated and borne out in the experiment that we conducted.