

Counterfactuals in Context

Dirk Buschbom

Dissertation
zur Erlangung des philosophischen Doktorgrades an der
Philosophischen Fakultät der Georg-August-Universität Göttingen

Göttingen 2014

Erstgutachterin: Prof. Dr. Regine Eckardt
Zweitgutachter: Prof. Dr. Manfred Sailer

Tag der mündlichen Prüfung: 03.06.2013

Acknowledgements

First of all, a very special thank you to Regine Eckardt, who was the main supervisor of this thesis. Next, I would like to thank also Manfred Sailer very much, who was the second supervisor. I am also grateful to Michael Job, Joost Kremers, and Edgar Onea for joining the defence committee, and to Wolfgang Hampe for handling all administrative affairs. A number of other people were involved at different stages of this dissertation. Thanks to Gert Webelhuth, Fritz Hamm, Anke Holler, Magda and Stefan Kaufmann, Götz Keydana, Heike Walker, Frank Veltman, and especially Jerra Lui Busch. In addition I would like to thank the audiences of the University of Göttingen's English Department's Linguistics Oberseminar, of Magda's Semantics Tea in Göttingen, and of the "Text Structures"-Courant Research Centre's LabMeeting in Göttingen. Thank you so much to Sarah Zobel for last minute discussion and incredible L^AT_EX-wizardry. Also Roland Schäfer and Niko Gerbl deserve credit for providing L^AT_EX-assistance. Last but not least, I would like to thank the Graduate School of Humanities Göttingen which supported the completion of this dissertation with a completion grant.

Contents

0	Introduction	1
1	The Data	5
1.1	Veltman's puzzle	5
1.1.1	More data – English	7
1.1.2	More data – German	9
1.2	A tentative account of the data in terms of the notion of negation-induced alternatives	10
1.3	Confirming the empirical reality of the concept of the negation-induced alternative: the particle <i>otherwise</i> as a discourse mirror	12
1.4	Generalizing to other alternative inducing data: focus and alternative questions	14
1.5	Extending the picture some more: problems for the sketch	17
1.6	Approaching a solution via the <i>because</i> -paraphrase of the data	21
1.6.1	<i>Because</i> and the NinA-challengers	23
1.7	Exhaustivity as a formal constraint on discourse felicity?	25
1.7.1	Wrapping things up	28
1.8	Summary	31
2	The Empirical Vicinity	34
2.1	Modal subordination	35
2.1.1	Preparing the ground: anaphora and discourse felicity	35
2.1.2	More data and a problem for the established constraints	38
2.1.3	The key to what is going on in the data and its relation to Veltman's puzzle	39

CONTENTS

2.1.4	Extending the empirical base, part 1: intensional contexts .	41
2.1.5	Preliminary summary	43
2.1.6	Extending the empirical base, part 2: modal subordination and presupposition	43
2.1.7	Recapitulating and connecting the observations: the gen- eral scheme	46
2.2	Contextually determined restrictions: principles and constraints . .	48
2.2.1	The semantic account in terms of entailments	49
2.2.2	Inadequacies of the semantic account – the pragmatic ap- proach and the notion of <i>accommodation</i>	53
2.3	Schueler (2008)	56
2.4	Summary	61
3	The Formal Framework	64
3.1	Locating Veltman’s concern	65
3.1.1	How complex is the process of making counterfactual as- sumptions?	65
3.1.2	The answer in a nutshell	66
3.2	Understanding counterfactuals: an informal motivation of their formal interpretation	66
3.2.1	The Ramsey-test and the shortcomings of a naive belief revision account	66
3.2.2	An argument against the Ramsey-test: Tichý 1	67
3.2.3	Tichý 1 as an argument against Stalnaker/Lewis	69
3.2.4	Interim conclusion	70
3.2.5	Tichý 2: more against Lewis’s position	71
3.3	The formalism	75
3.3.1	Some definitions of elementary concepts: <i>update</i> , <i>sup-</i> <i>port</i> , and <i>entailment</i>	75
3.3.2	The system’s ontology	77
3.3.3	The kernel of Veltman’s update semantics	79
3.3.4	What it means to make counterfactual assumptions: pre- liminaries	84

CONTENTS

3.3.5	What it means to make counterfactual assumptions: the formal account	86
3.4	Excursion: putting theory to practice – computing Tichý’s examples in Veltman’s system	89
3.4.1	Tichý 1	93
3.4.2	Tichý 2	103
3.5	Summary	110
4	The Analysis	112
4.1	The procedural side of discourse computation	113
4.1.1	The communicative environment of the discourse in the initial state	114
4.1.2	Updating with a NinA-trigger: the discourse initial sentence <i>John did not drink</i>	115
4.1.3	Communicating <i>He would have become sick</i>	116
4.1.4	Interpreting <i>He would have become sick</i> – a naive first attempt	118
4.1.5	The revised full account	119
4.2	The Semantic account	121
4.2.1	The first case: elliptic discourse with nomological presupposition	121
4.2.2	The second case: radical interpretation of counterfactuals – non-elliptic discourse without nomological presuppositions	123
4.2.3	The third case: the infelicitous # <i>John drank. He would not have become sick</i>	126
4.2.4	The fourth case: <i>John drank. If he had not drunk, he would not have become sick</i>	129
4.2.5	Generalizing the point	138
4.3	The solution to Veltman’s puzzle: felicity conditions for elliptic counterfactuals	138
4.4	On the semantic equivalence of counterfactual discourse and its <i>because</i> -paraphrase	142

CONTENTS

4.4.1	First observations regarding the synonymy of the data . .	143
4.4.2	The semantics of <i>because</i>	146
4.4.3	On the presupposition of <i>because</i> and the semantic relation between elliptic counterfactual discourse and the <i>because</i> -paraphrase.	150
4.4.4	Aside: why we all are counterfactual drunkards – the tenability of the presupposition of <i>because</i>	152
5	Conclusion	155
	Bibliography	169

Chapter 0

Introduction

This thesis is concerned with a certain kind of counterfactual conditional, namely the elliptic variant occurring in discourses like e.g.

(1) John did not drink any wine. He would have become sick.

Frank Veltman makes in [Veltman \(2005\)](#) as an aside a rather intriguing observation¹: while (1) paraphrases the full conditional

(2) If John had drunk any wine, he would have become sick.

the negative counterpart of (2) does not allow for a paraphrase like (1) in which the antecedent of the conditional is left implicit. Even though – just as in the above examples – the first sentence of (4) corresponds to the presupposition of the *if*-clause of the full counterfactual in (3), the resulting sequence of sentences is pragmatically not acceptable, but infelicitous (and thus marked by #):

(3) If John had not drunk too much wine, he would not have become sick.

(4) # John drank too much wine. He would not have become sick.

This strange asymmetry in discourse behavior and paraphrasability of counterfactual *if*-sentences – to which I will refer to as *Veltman's puzzle* – provides the starting point and initial concern of this thesis. In the following analysis I will

¹Cf. [Veltman \(2005\)](#), p. 160.

CHAPTER 0. INTRODUCTION

formulate a solution to this problem that will go beyond the narrow confines of the issue, but will be of general importance for understanding the semantics and pragmatics of counterfactuals.

The analysis divides into four chapters. Chapters 1 and 2 are concerned with informal empirical groundwork, while Chapter 3 and Chapter 4 are devoted to the formal aspects of the analysis.

In Chapter 1 an in depth introduction to the data is delivered in order to get acquainted with the problem and to sharpen intuitions. Based on observations that will be spelled out in the course of the chapter the phenomenon and research questions are more neatly defined. By focussing on the role that negation seems to play for the felicity of the data also a first working hypothesis is formulated. It suggests that discourse felicity is determined by the presence resp. absence of a *negation induced alternative* – a concept for which clear and independent empirical confirmation will be delivered – that acts as a contextually determined filler for the implicit antecedent. Following this lead the concern is generalized and other alternative inducing phenomena like *focus* are briefly discussed as possible discourse contexts that allow to leave the antecedent of a counterfactual conditional implicit. However, a set of additional examples that clearly defy the sketched proposal necessitates a re-assessment of the phenomenon and of the initial working hypothesis. The key to the successful analysis is eventually gained after some further contemplation upon the semantics of the discourses of concern here. Crucial is the observation that for most of the data under consideration a paraphrase in form of a *because*-sentence is available.

After having gained a clearer idea of the phenomenon and how to account for it in Chapter 1, the informal overview is rounded up in Chapter 2 by a discussion of the empirical (and theoretical) vicinity of the phenomenon. The concern lies with data that follows a similar scheme as the discourses I am concerned with in that contextually supplied implicit information plays an important role in its interpretation and felicity. First, the concept of *modal subordination* (cf. for instance Roberts (1987), Roberts (1989), and Roberts (1996a)) is discussed. Second, the

CHAPTER 0. INTRODUCTION

elliptic counterfactuals examined in Schueler (2008) are addressed. Although on the surface similar, it turns out that the data and analysis with which Schueler is concerned are incompatible with what is at issue here. With respect to modal subordination more overlap exists. The discussion however reveals that either – contrary to what Roberts suggests – the discourses under consideration can not be subsumed under the phenomenon, or that the concept of modal subordination is too narrowly defined.

In Chapter 3 the formal framework in which the analysis will be spelled out is introduced. Point of reference here is the update semantics for counterfactuals formulated in Veltman (2005). Veltman (2005) thus provides not just the starting point for the empirical aspect of this thesis, but also for its theoretical account. As a dynamic framework Veltman’s update theory takes the essence of meaning to be its potential to change information states. In the first half of this chapter Veltman’s particular account of counterfactuals finds detailed motivation, while in the second half the formal definitions are presented and expounded.

Finally, in Chapter 4 the solution to Veltman’s puzzle is laid out. Elaborating on the idea that the logical structure of the premise set is crucial for the explanation of the data – which gained contour at the end of Chapter 1 in connection with the observation that a *because*-paraphrase exists for the examples – and employing the formal tools presented in Chapter 3, a semantic account for the pragmatic infelicity of the discourses under consideration is given. Concomitant to this comes another result of general importance that also has, as far as I know, not been addressed in the literature. Based on the observation that counterfactuals without appropriate contextual support are more often than not accepted as true, although they are trivially false, a pragmatic account of the semantic acceptability of these cases is developed. Adapting a notion introduced by Donald Davidson in Davidson (1973), the analysis is about the *radical interpretation of counterfactuals*. Finally, also questions concerning the semantics of the *because*-paraphrase from which the crucial insight for the solution of Veltman’s problem derives are addressed. Doing so a revision of a Lewis-style semantics of indicative *because* (cf. Lewis (1973a)) is formulated that captures the meaning of subjunctive *because*-sentences, and

CHAPTER 0. INTRODUCTION

a certain issue connected to the presuppositions of *because* is analyzed and accounted for.

Chapter 1

The Data

1.1 Veltman's puzzle

In an effort to motivate a decompositional semantics for counterfactual conditionals Frank Veltman observes in Veltman (2005) that the antecedent in a counterfactual figures as only but one possible (syntactic and semantic) context for the *would have* + past participle verb phrase projection that acts as the conditional's consequent. To illustrate the point Veltman provides the following examples (cf. Veltman (2005), p. 160):

- (1) If John had drunk any wine, he would have become sick.
- (2) John did not drink any wine. He would have become sick.

Not only does the consequent of (1) reappear as the second sentence of (2), but moreover, both (1) and (2) seem to be synonymous. (At least on the account that a counterfactual antecedent like *If John had drunk any wine* presupposes the factual truth of the sentence *John did not drink any wine*.¹) While, as Veltman points out, "[s]entences with a verb phrase consisting of 'would have' + past participle make no sense if they are presented without context" (Veltman (2005), p. 160),

¹Although this presupposition does not generally apply (for counterexamples see e.g. Anderson (1951)) the data to be discussed will exclusively be data with a false (i.e. presuppositional) antecedent. Most of the examples bear contexts that entail the truth of the complement proposition to the counterfactual's antecedent.

CHAPTER 1. THE DATA

the sentence *He would have become sick* may nonetheless arguably be treated as a separate update² in the mini-text (2), since it appears here completely unmediated by any linguistic means. Based on this observation Veltman suggests to treat the complex of *would have* + past participle also as a separate update when it occurs as the consequent of a conditional.

While so far the story is rather trifling, the case becomes interesting, and indeed problematic, if one tries to generalize the idea. A slight revision of the data illustrates what is at stake. Consider the following data given by Veltman (Veltman (2005), p. 160):

- (3) If John had not drunk too much wine, he would not have become sick.
- (4) # John drank too much wine. He would not have become sick.

The problem – *Veltman's puzzle* – that arises from these two examples becomes clear when contrasted with the pair of examples given in (1) and (2) : just as in the first set of data the second sentence of (4) reappears as the consequent in the conditional (3). But although the antecedent of (3) and the initial sentence in (4) follow exactly the pattern of fact/counterfact exhibited in (1) and (2), the mini-text (4) is not an acceptable sequence of sentences (which is therefore marked by '#').³

While Veltman proposes an elaborate semantics for counterfactuals for which the data just given act only as an opener, he does not discuss the asymmetry in the examples in detail, but gives only a rudimentary sketch.⁴ However, once the data base is extended, it turns out that the problem is rather intricate and that its analysis allows for an instructive perspective on the interplay between the semantics and pragmatics of counterfactuals and the construction of elliptic discourse.

²The concept of an *update* in dynamic semantics is neatly captured in the following quote: “On the dynamic view, knowing the meaning of a sentence is knowing the change it brings about in the cognitive state of anyone who wants to incorporate the information conveyed by it.” (Veltman (2005), p. 160) The ‘change brought about’ is eventually defined as an *update*.

³The convention to mark the *infelicity* of data by a prefixed ‘#’ will be maintained throughout the whole study. Veltman uses ‘*’ instead.

⁴For a relevant quote from Veltman (2005), see my footnote 9, page 11.

CHAPTER 1. THE DATA

Outline of the chapter

In the following sections of this chapter the empirical side of the described problem shall be further explored. After extending the data base at hand of more English and German examples in sections 1.1.1 and 1.1.2 the concept of the *negation-induced alternatives* is introduced in section 1.2 as a promising key to the solution of the problem. Subsequently in section 1.3 the empirical reality of the concept of the negation-induced alternatives is confirmed. Based on the informal proposal sketched in 1.2 and 1.3 a generalization to data containing other alternative inducing elements like focus resp. alternative questions is attempted in section 1.4, but first doubts with respect to the feasibility of the lead arise. These concerns are eventually deepened when in section 1.5 more data are introduced that clearly defy the proposal to analyze Veltman's puzzle in terms of the concept of the negation-induced alternatives. A more promising approach gains however contour when in sections 1.6 and 1.7 finally the possibility of a *because*-paraphrase for Veltman's examples is studied. The observations formulated here will ultimately provide with the key to solving the problem.

1.1.1 More data – English

Before proposing a first informal account of the problem that will be based on what I call *negation-induced alternatives*, more examples shall be presented that give an impression of the extent and systematics of the phenomenon.

In the following mini-discourses the initial sentence factually describes John's drinking behaviour, while the subsequent counterfactual describes what would have been the case if John's behaviour had been opposite to his actual behaviour.

- (5) John did not drink any wine. If John had drunk any wine, he would have become sick.
- (6) John drank too much wine. If John had not drunk too much wine, he would not have become sick.

If one skips the antecedent of the conditional, the already encountered aberration

CHAPTER 1. THE DATA

may be observed again. While (7) below is inconspicuous, (8) is infelicitous and indeed not interpretable:

- (7) John did not drink any wine. He would have become sick.
- (8) # John drank too much wine. He would not have become sick.

Note that this pattern seems to be rather general, as the following mini-discourses show. Again, the initial sentence factually details a certain scenario, while the subsequent counterfactual describes what would have been the case if the initial situation had been its opposite.

- (9) The glass was not dropped. If it had been dropped, it would have broken.
- (10) The glass was dropped. If it had not been dropped, it would not have been broken.

When the antecedent is dropped, one example remains unproblematic, while the other turns out infelicitous and incoherent:

- (11) The glass was not dropped. It would have broken.
- (12) # The glass was dropped. It would not have been broken.

The same behaviour recurs when switching from the above causal scenarios to a deontic context in which obligations and permissions provide the frame of reference:

- (13) John did not steal the purse of the old lady. If John had stolen the purse of the old lady, he would have been punished.
- (14) John did not steal the purse of the old lady. He would have been punished.
- (15) John stole the purse of the old lady. If John had not stolen the purse of the old lady, he would not have been punished.
- (16) # John stole the purse of the old lady. He would not have been punished.

Also for bouletic scenarios (i.e. scenarios in which desires and wishes are of concern) the encountered asymmetry in felicity can be observed:

CHAPTER 1. THE DATA

- (17) John did not steal the purse of the old lady. If John had stolen the purse of the old lady, his father would have been sad.
- (18) John did not steal the purse of the old lady. His father would have been sad.
- (19) John stole the purse of the old lady. If John had not stolen the purse of the old lady, his father would not have been sad.
- (20) #John stole the purse of the old lady. His father would not have been sad.

1.1.2 More data – German

The phenomenon occurs also cross-linguistically, as a translation of the English data into German shows. Example (21) translates example (5), (22) translates (7), (23) translates (6), and (24) translates (8):⁵

- (21) Hans trank keinen Wein. Wenn Hans Wein getrunken
Hans drink_{+past} NEG wine. If Hans wine drink_{pp}
hätte, wäre ihm übel geworden.
AUX_{+past+sbjv}, AUX_{+past+sbjv} PRO_{3.ps+sg+m+dat} sick become_{pp}.
- (22) Hans trank keinen Wein. Ihm wäre übel
Hans drink_{+past} Neg wine. PRO_{3.ps+sg+m+dat} AUX_{+past+sbjv} sick
geworden.
become_{pp}.
- (23) Hans trank zuviel Wein. Wenn Hans nicht zuviel Wein
Hans drink_{+past} too_much wine. If Hans NEG too_much wine
getrunken hätte, wäre ihm nicht übel
drink_{pp} AUX_{+past+sbjv}, AUX_{+past+sbjv} PRO_{3.ps+sg+m+dat} NEG sick
geworden.
become_{pp}.

⁵The abbreviations used in the glosses are: *AUX* = auxiliary, *NEG* = negation, *PRO* = pronoun, *+past* = tense: past, *pp* = past participle, *+sbjv* = mood: subjunctive, *3.ps* = person: 3rd, *+sg* = number: singular, *+m* = gender: masculine, *+dat* = case : dative. Only dative pronouns are marked for case.

CHAPTER 1. THE DATA

- (24) # Hans trank zuviel Wein. Ihm wäre nicht
Hans drink_{+past} too_much wine. PRO_{3.ps+sg+m+dat} AUX_{+past+sbjv} NEG
übel geworden.
sick become_{pp}.

Since in the overwhelming majority of examples the German translation behaves just as the English data, I will provide further German data and translations from English to German only whenever the picture is either to be completed, or when deviations in intuitions are to be highlighted.

1.2 A tentative account of the data in terms of the notion of negation-induced alternatives

In the analysis to be proposed I do not follow Frank Veltman's suggestion to treat the *would have* + past participle VP-projection as a largely autonomous unit that could occur outside of a conditional context.⁶ The complex is rather considered to appear as the consequent of an *implicit antecedent* in the data above. Under this point of view (2) actually reads as *John did not drink any wine. [If he had drunk any wine] he would have become sick.*⁷ Answering the question where the hidden antecedent comes from will also provide with a first hypothesis about why e.g. examples (4), (8), and (12) are incoherent. I will propose that the hidden antecedent is given as a procedural or cognitive trade-off connected to the occurrence of negation in the preceding context: correlated to e.g. the initial sentence of (2), *John did not drink any wine*, is the supposition that drinking is within the behavioural capabilities of John. The intuition that negation induces such an affirmative alternative is arguably irrevocable. Even in cases like *John did not drink the ocean dry* or *John did not light the sun* the intuition is that it is – absurd as it may be – within the agent's action alternatives (i.e. possibilities) that John could literally drink the ocean dry, etc.⁸ This happenstance becomes relevant in

⁶For a stance similar to Veltman's see for instance Kasper (1987).

⁷While at first sight this may seem to undermine Veltman's proposal to treat the semantics of conditionals in terms of two separate updates, the point is nonetheless actually rather superficial and does not affect the substance of his semantics in any way.

⁸This point seems to fall in line with findings concerning the processing and representation of negated information. See for instance Kaup et al. (2006), and Kaup et al. (2007) (but see also Tian

CHAPTER 1. THE DATA

combination with counterfactual conditionals: while counterfactuals describe alternative scenarios, negation induces a potential, indeed counterfactual alternative that could link up to the scenario given by a counterfactual.

The generalization that may be formulated on this ground is thus that the negation-induced alternative of a negated, discourse-initial sentence may serve as the antecedent of a subsequent counterfactual consequent whenever the discourse describes a fact/counterfact-scenario. Indeed, in a discourse sequence of an initial negated sentence and a subsequent counterfactual consequent that lacks an overt antecedent the negation-induced alternative (in the following abbreviated as *NinA*) of the initial sentence is a natural candidate to figure as an implicit, contextually determined antecedent. Negative discourse initial sentences thus seem to provide a proper context for the *would have*-modal perfect in the consequent. The data arguably confirm this thesis: unlike in (2), (7), (11) in the examples (4), (8), (12) no such suppositional contextual fillers are available, since here the preceding context is affirmative and therefore does not induce an alternative. Since however the modal perfect is interpretable only within a conditional context, the lack of an antecedent leaves the counterfactual incomplete. Hence, (4) is rendered unacceptable. And similar for (8) and (12).⁹

The same line of reasoning could also explain the following data that at first sight

et al. (2010) for a critical evaluation of Kaup et al.'s work).

⁹ Although Veltman differs in his assessment of the role of the counterfactual consequent, the first account formulated here seems nonetheless compatible with what he thinks about the role of negation in the initial sentence. Cf. Veltman (2005:p. 161), where he writes: "Apparently, after processing the first sentence of (i) [= example (2) here, D. B.] the stage is set for the interpretation of the subsequent sentence, whereas in (iii) [= example (4) here, D. B.] this is not the case. When you have to interpret a negative sentence, such as the first sentence in (i) [= (2), D. B.], the interpretation process starts with an update with the positive subsentence, and then continues with some operation on this intermediate result. This intermediate result is kept in memory as an auxiliary state subordinate to the main state, and differing from it mainly in that it supports a statement that is rejected in the main state. This subordinate state is the state the modal perfect in the second sentence of (i) [= (2), D. B.] is looking for. In interpreting the first sentence of (iii) [= (4), D. B.] no such subordinate state is created. Therefore the second sentence of (iii) [= (4), D. B.] finds nothing to pertain to." Other conceptualizations of this kind of phenomenon can be found in Kaufmann (2000), and van Rooy (2005), as Veltman points out. But also Roberts (1987) and Frank (1997) formally address the issue (with reference to the account of negation given in Discourse Representation Theory (cf. Kamp (1981) and Kamp and Reyle (1993))).

CHAPTER 1. THE DATA

seems to rebut the proposal for exhibiting the ‘wrong’ distribution of negation:

- (25) John stayed sober. [*antecedent to be filled*] He would have become sick.
(26) # John did not stay sober. [*antecedent to be filled*] He would not have become sick.

One might argue – admittedly somewhat tentatively, though – that the reason why (25) is an acceptable sequence, although the initial sentence lacks an overt negative marker, is that it contains a lexicalized negation. On the assumption that *to stay sober* is synonymous to *to not drink*, the available action alternative to this is the affirmative *to drink*. While thus a filler for the empty antecedent is conceivable that describes a scenario in which no background laws are violated, no such constellation can be received in case of (26). What appears as a single negation in the first sentence of (26) is indeed a double negation with *to not stay sober* actually meaning the affirmative (*not not*) *to drink*. Hence, *to drink* is the only information that could enter the empty antecedent. But this would create a counterfactual scenario that violates our knowledge about the effects drinking might have. *If he had drunk he would not have become sick* is certainly not the ‘normal’ situation. Thus, (26) is rendered incoherent.

1.3 Confirming the empirical reality of the concept of the negation-induced alternative: the particle *otherwise* as a discourse mirror

Note that the analysis of another phenomenon that Veltman brings to attention in connection with the data under discussion eventually supports the conception of the negation-induced alternative. With respect to example (4) Veltman observes in passing that the insertion of the particle *otherwise* into the discourse will alter the picture completely (cf. Veltman (2005), p. 160).¹⁰ Adding *otherwise* to the

¹⁰A similar observation with respect to the semantic impact of the particle *otherwise* is made in Frank and Kamp (1997) on p. 164.

CHAPTER 1. THE DATA

infelicitous (4) (repeated here as (27)) renders the mini-discourse (28) not only fully acceptable, but indeed synonymous to (3), hence mimicking the behaviour of (1) and (2).

(27) # John drank too much wine. He would not have become sick.

(28) John drank too much wine. Otherwise he would not have become sick.

While thus the observed asymmetry in discourse behaviour is revoked, it is interesting to note that the particle *otherwise* at first sight may seem to be semantically empty, though. This is suggested by the insertion of *otherwise* into (29). There is no manifest difference in meaning between (29) and (30) (or in any of the other felicitous examples for this matter if *otherwise* is inserted there):

(29) John did not drink any wine. He would have become sick.

(30) John did not drink any wine. Otherwise he would have become sick.

The strange behavior of the data becomes explicable if one assumes that *otherwise* is synonymous to *IF NOT* $\phi_{\text{counterfactually}}$, with the negation picking up and operating on the preceding propositional context ϕ .¹¹ In case of (28) with the antecedent to the *would have*-complex in the consequent of the conditional thus spelled out one arrives at *IF NOT (John drank any wine)_{counterfactually} he would not have become sick* – which is perfectly ok, while *otherwise* in (30) yields the conditional *IF NOT (John did not drink any wine)_{counterfactually} he would have become sick*. Since the ensuing double negation in the antecedent is equivalent to the NinA (*If*) *John had drunk any wine* one arrives at the exact same (hidden) antecedent that the negation in the initial sentence of (29) induces. Hence the impression that the insertion of the particle *otherwise* in the sequence is semantically vacuous. While in effect *otherwise* works like a semantic ‘discourse mirror’ – it picks up and reverses the initial (propositional) context and thus is able to fill the empty antecedent with the required information – the confrontation of examples (29) and (30) at the same time positively confirms the empirical reality of the notion of the negation-induced

¹¹The question about what scope *otherwise* takes is rather intriguing, but goes beyond the present concern and will thus not be addressed in any detail. For a first study see Fuhrmann (2011).

CHAPTER 1. THE DATA

alternative: only because both examples share the same implicit antecedent can the impression arise that the particle *otherwise* is semantically void. If one follows intuition and assumes that the expression *otherwise* indeed figures as a discourse mirror the concept of the negation-induced alternatives provides a natural and compelling concept to account for the equivalence of (29) and (30).

Summary: the notion of the NinA and its empirical reality

On basis of the notion of the negation-induced alternative the proposed analysis is able to account for certain phenomena concerning the discourse behaviour of counterfactual conditionals. In particular it is able to spell out in detail under what conditions the antecedent of a counterfactual conditional may remain implicit in discourse. It counters the assumption that the modal perfect that acts as the consequent of a counterfactual conditional may appear outside of the context of a preceding *if*-clause, and explains the effect that the insertion of the particle *otherwise* into the data has.

In summary the picture with respect to the notion of the negation-induced alternative is this: while negated sentences induce (trivially counterfactual, positive) alternatives, the isolated *would have*-complex stands as the consequent of an incomplete conditional. It is waiting for an appropriate (counterfactual) antecedent. In sequence the initial negation and the isolated *would have*-remnant are thus a perfect, complementary match: the former provides in form of the negation-induced alternative the counterfactual antecedent that the latter is lacking.

1.4 Generalizing to other alternative inducing data: focus and alternative questions

Before turning to some challenges to the story told so far it first shall be noted that the proposed account of discourse construction involving contextually given information can apparently be generalized to other alternative inducing phenom-

CHAPTER 1. THE DATA

ena, like e.g. focus and alternative questions.¹² Since both provide with (implicit, respectively explicit) alternatives one may predict that both constitute suitable contexts for the modal perfect. The next set of data suggests that this seems to be indeed the case. The fact that a counterfactual consequent may link up to these in discourse – as demonstrated in the below examples (in which the actually implicit antecedent is made explicit) – arguably delivers a parallel model for and independent support of the account in terms of negation induced alternatives given so far.¹³ –The question of course being what other linguistic phenomena evoke alternatives, and whether these also provide suitable contexts for the modal perfect. Just as in the case of the negation-induced alternatives the alternatives to the focussed constituent resp. the answer alternatives mentioned in the alternative question remain available in the context of the discourse (more precisely, in the counterfactual dimension of the conversational background). Eventually they serve to fill in the implicit counterfactual antecedent (for perspicuity given here in brackets) required by the *would have*-consequent.

- (31) [MARY]_F got the job. (If John had gotten it) John would have made a mess out of it.
- (32) Q: Would you like tea or coffee? A: Tea. (If I had coffee) Coffee would get me too agitated.

¹²For both of these topics a vast body of semantic literature exists. I will however rely on an intuitive understanding of both notions only, since the point shall not be studied in any detail. A classical reference for an alternative semantics for focus is Rooth (1985), formal semantics for alternative questions are given for instance by Groenendijk and Stokhof (1984), Biezma and Rawlins (2011), Romero and Han (2003).

¹³Although Veltman does not seem to follow the generalization (or the idea that negation may induce alternatives in its own right), he also points out that focus generates suitable contexts for the modal perfect. In footnote 4 of Veltman (2005) on p. 160 he writes: “For readers who think that it is crucial that there be a negation in the first sentence, it will be worthwhile to look at [...] *We asked Mary to taste the wine. John would have become sick.* If ‘Mary’ is stressed in the first sentence and ‘John’ in the second, the second sentence makes perfect sense.” Veltman continues to elaborate on this example in footnote 6, p. 161 in Veltman (2005): “To deal with the example of footnote 4 one needs to invoke a theory of focus like the one presented in Rooth (1985). According to this theory the general function of focus is evoking alternatives. (In this case, alternatives to Mary –which other people could we have asked to taste the wine?). (sic.) These alternatives will give rise to subordinate states and are available for sentences in which the modal perfect is used to be interpreted in. (Apparently, we are ready to accommodate the idea that John is one of the alternatives.)”

CHAPTER 1. THE DATA

The point is supported by the fact that (31) and (32) each have an *otherwise*-synonym:

- (33) [MARY]_F got the job. Otherwise, John would have made a mess out of it.
- (34) Q: Would you like tea or coffee? A: Tea. Otherwise I would get too agitated.¹⁴

Thus, the examples suggest a certain commensurability between the data and seem to support the informal account sketched so far. Although the issue shall not be explored in any depth, some more observations shall nonetheless at least be mentioned. First, note that while the antecedent may in both cases remain implicit, the alternative in the focus-examples apparently needs to be made explicit in the consequent.¹⁵

- (35) Did Mary or John get the job? Mary got the job. We would have been sad.

This can not mean – unlike when the particle *otherwise* is inserted into the discourse – that *If John had gotten the job, we would have been sad*.

Although these observations are interesting in their own right, they shall not be discussed any further. More important seems however the next data to follow. The examples anticipate a problem that will be of concern later on. Since focus induces alternatives in all contexts, one may hypothesize that the occurrence of negation in the initial sentence is redundant. I.e. the prediction would be that the discourse below in (37) would be acceptable in any case:

- (36) [MARY]_F did not get the job. If John had not gotten the job, he could not have made a mess out of it.

¹⁴The contrastive focus that one might expect is suppressed here (and also in example (35)). It is irrelevant for the discussion.

¹⁵Which eventually casts doubt on the necessity of a marking via an intonational focus. Without an explicit mentioning of the alternative the data apparently become infelicitous (the reason being arguably that without a salient alternative it is not clear which of the alternatives could figure as the implicit antecedent).

CHAPTER 1. THE DATA

- (37) # [MARY]_F did not get the job. John could not have made a mess out of it.

Although the discourse context is enriched not just by the NinA, but also by the focus alternatives, (37) is nonetheless infelicitous. An extra twist to this is added when inserting the particle *otherwise*. This again renders the discourse acceptable:

- (38) [MARY]_F did not get the job. Otherwise John could not have made a mess out of it.

Apparently, there is more to the story of the contextually determined implicit alternative as a mediator of discourse felicity than hypothesized so far. I will later on propose a solution to the issue connected to (38), but first some more problems shall be highlighted in the next section.

1.5 Extending the picture some more: problems for the sketch

While the story of the negation-induced alternatives seems rather potent in accounting for the data, it does not come without problems. Indeed, there are examples that apparently defy or even contradict the proposal. The following collection of mini-discourses demonstrates the point. One discourse challenges the account insofar as it matches the identified NinA-pattern, but nonetheless fails the story: even though the initial sentence contains a negation and thus arguably provides a negation-induced alternative the discourse is *not* acceptable. Another example escalates the challenge even further. The discourse does not just defy the proposal, but moreover contradicts it insofar the discourse is felicitous although the context fails to supply with an implicit antecedent. The confrontation with these examples will eventually lead to a new assessment of the phenomenon and the issues deriving from it and will necessitate a re-evaluation of the account given so far.

Keeping this in mind consider now the following mini-texts with an initial negative sentence and a subsequent counterfactual conditional:

CHAPTER 1. THE DATA

(39) John did not have surgery. If he had had surgery, he would have died.

Dropping the antecedent of the counterfactual generates a discourse that is synonymous to the counterfactual that is part of (39):

(40) John did not have surgery. He would have died.

Varying slightly the scenario (by negating the consequent) such that a counterfactual situation with reverse consequences is described, acceptability changes. Dropping the antecedent of the counterfactual generates in this case an infelicitous discourse even though the NinA-account and Veltman's own conception (as quoted in footnote 9) predict its felicity:

(41) # John did not have surgery. He would not have died.

Next, consider the case in which the discourse initial sentence is not negative:

(42) John had surgery. If he had not had surgery, he would not have died.

When this time the antecedent is dropped, the resulting discourse is, as predicted, infelicitous:

(43) # John had surgery. He would not have died.

However, when also the negation in the consequent is dropped, the resulting sequence all of a sudden becomes acceptable. Indeed, (44) is synonymous to the counterfactual in (45):

(44) John had surgery. He would have died.

(45) John had surgery. If he had not had surgery, he would have died.

This of course is rather remarkable, since the proposal of the negation induced alternatives would predict the infelicity of (44), since the affirmative initial sentence can not enrich the context with a NinA to which the modal perfect of the counterfactual consequent could link up to to complete the counterfactual conditional.

CHAPTER 1. THE DATA

What is further worth to be noted is that the insertion of the particle *otherwise* into the discourse again renders the infelicitous examples felicitous, while the felicity and meaning of the the acceptable discourses remains unaffected by it (with (48) and (49) being completely synonymous to (40) and (44)):

- (46) John did not have surgery. Otherwise he would not have died.
- (47) John had surgery. Otherwise he would not have died.
- (48) John did not have surgery. Otherwise he would have died.
- (49) John had surgery. Otherwise he would have died.

The same phenomenon can be observed also in the German translation of the data. (Example (50) below translates example (41) above, and the same for the pairs (51) and (40), (52) and (43), and (53) and (44).¹⁶) A negated counterfactual consequent fails also here to present a felicitous continuation not just of a positive initial sentence (which does not supply a NinA), but also of its negative counterpart (which does supply a NinA):

- (50) # Hans wurde nicht operiert. Er wäre nicht gestorben.
Hans AUX_{psv+past} NEG operate_on. He AUX_{+past+sbjv} NEG die_{pp}.
- (51) Hans wurde nicht operiert. Er wäre gestorben.
Hans AUX_{psv+past} NEG operate_on. He AUX_{+past+sbjv} die_{pp}.
- (52) # Hans wurde operiert. Er wäre nicht gestorben.
Hans AUX_{psv+past} operate_on. He AUX_{+past+sbjv} NEG die_{pp}.
- (53) Hans wurde operiert. Er wäre gestorben.
Hans AUX_{psv+past} operate_on. He AUX_{+past+sbjv} die_{pp}.

The insertion of discourse mirrors like *sonst* or *andernfalls* into the discourse once more changes the picture. They render the initially infelicitous data acceptable¹⁷:

¹⁶The impersonal passive that occurs in the initial sentence of the examples is only detailed by the introduction of the category *AUX_{psv}* that designates the passive auxiliary; all other abbreviations are as before.

¹⁷(54) translates (46), while (55) translates (47).

CHAPTER 1. THE DATA

- (54) Hans wurde nicht operiert. Andernfalls/Sonst wäre er
 Hans AUX_{psv+past} NEG operate_on. Otherwise AUX_{+past+sbjv} he
 nicht gestorben.
 NEG die_{pp}.
- (55) Hans wurde operiert. Andernfalls/Sonst wäre er nicht
 Hans AUX_{psv+past} operate_on. Otherwise AUX_{+past+sbjv} he NEG
 gestorben.
 die_{pp}.

Focussing on the English examples without the particle *otherwise* the observations are summarized in the following table:

example nr.	initial sentence	antecedent	consequent	felicity
(40)	negative	NinA	positive	√
(41)	negative	NinA	negative	infelicitous
(43)	positive	none	negative	infelicitous
(44)	positive	none	positive	√

(Cross)comparing the discourses it suggests itself to reconsider the proposal. Arguably, the key to discourse felicity in the examples is not simply the (undisputable) availability of a NinA, but in fact the occurrence of (sentential) negation in the consequent. While this does not contradict the existence of the concept of the negation-induced alternative as a whole, these considerations however lead to question the role of the notion in explaining the behaviour of the data. If it is not just the availability of a NinA, but apparently the occurrence of negation in the counterfactual consequent that determines discourse felicity, the question is of course how this can be accounted for. A first analysis of this issue will be attempted in due course.

1.6 Approaching a solution via the *because*-paraphrase of the data

From the above it should be obvious that what initially seemed to be clear has turned out to be rather foggy. There is a need for a re-assessment of the data and of the naive analysis. A first step in doing so is provided by a simple diagnostics that involves a commonplace paraphrase of the discourses in question in terms of a *because*-sentence. This will expand the picture in a promising way. The initial example (here repeated as (56)) suits to make the case in point. As it stands, in the particular constellation of factual and counterfactual statement, the mini-discourse

(56) John did not drink any wine. He would have become sick.

may be understood to be synonymous to

(57) Because John would have become sick, he did not drink any wine.¹⁸

This scheme works for all the unproblematic data given so far. This holds not just for those examples that express a clear causal nexus like in (56) and (58), but also for the discourses involving a deontic scenario (here repeated as (60)), and the bouletic variant given earlier (which is repeated here as (62)):

(58) The glass was not dropped. It would have broken.

(59) Because it would have broken, the glass was not dropped.

(60) John did not steal the purse of the old lady. He would have been punished.

(61) Because he would have been punished, John did not steal the purse of the old lady.

(62) John did not steal the purse of the old lady. His father would have been sad.

(63) Because his father would have been sad, John did not steal the purse of

¹⁸One has to be cautious here with the word ‘synonymous’, since the causal flavor that characterizes the particular example need not always be present. However, for most of the data under discussion this is a dominant trait.

CHAPTER 1. THE DATA

the old lady.

The same holds also for the German translations of the examples. The German variant of the initial discourse shall suffice to illustrate the point:

- (64) Hans drank keinen Wein. Ihm wäre übel
 Hans drink_{+past} Neg wine. PRO_{3.ps+sg+m+dat} AUX_{+past+sbjv} sick
 geworden.
 become_{pp}.
Hans did not drink any wine. He would have become sick.
- (65) Weil ihm übel geworden wäre, trank Hans
 Because PRO_{3.ps+sg+m+dat} sick become_{pp} AUX_{+past+sbjv}, drink_{+past} Hans
 keinen Wein.
 Neg wine.
Because he would have become sick, Hans did not drink any wine.

With respect to the infelicitous discourses it is important to note that also their *because*-analogues are unacceptable. The paraphrase has thus not just the same truth conditions, but also the same felicity conditions.¹⁹

- (66) #John drank too much. He would not have become sick.
 (67) #Because he would not have become sick, John drank too much.
 (68) #The glass was dropped on a hard surface. It would not have broken.
 (69) #Because it would not have broken, the glass dropped on a hard surface.²⁰

The fact that the *because*-sentences show the exact same behaviour as the respective discourses confirms the significance of the paraphrase. What renders the point even more interesting is (i) the availability of the *because*-paraphrase also in connection with the other alternatives invoking data -the focus and alternative question examples- mentioned already, and (ii) the behaviour of the *because*-analogue for those cases that challenge the NinA-story. While the former shall only be mentioned without going into detail, the latter will be elaborated on more extensively.

¹⁹This suggests to assume that the connective *because* lexicalizes the specific discourse relation that is so characteristic of the NinA-examples.

²⁰That for data of this kind a secondary, maybe marginal, but felicitous reading (in terms of an *in order to*-interpretation) seems to be available will be addressed later on.

CHAPTER 1. THE DATA

The *because*-variants of the focus and alternative question examples are:

- (70) Because John would have made a mess out of it, [MARY]_F got the job.²¹
(71) Because coffee would get me too agitated, I would like tea.

Notwithstanding the problems already mentioned, the availability and consistency of the *because*-paraphrase also for the focus and alternative question-examples is – together with the data’s behavior in connection with the particle *otherwise* – rather remarkable.²²

1.6.1 *Because* and the NinA-challengers

To begin the discussion of the *because*-paraphrases in relation to those examples that challenge the NinA-account I will return to the *surgery*-data introduced above in (40), (41), (43), and (44). The main point derives from the contrast between (41) and (44). Recall that the example in (41) is infelicitous, although the NinA-account would predict it to be felicitous: while the discourse initial sentence provides a NinA to fill in the implicit antecedent, the sequence is still not acceptable. (44) however is felicitous, whereas according to the first account sketched

²¹Although just a minor point, note that the *because*-paraphrase of the focus example is somewhat more restrictive in its meaning than the corresponding discourse. As it stands the *because*-sentence gives a mono-causal reason for Mary getting the job, while John did not get it. This kind of determinacy is absent in the original example (31): [MARY]_F got the job. John would have made a mess out of it.

²²The question whether every discourse has a *because*-paraphrase and whether every *because*-sentence can be expressed in a corresponding elliptic counterfactual discourse shall not be discussed in detail. Just note that not every discourse has a synonymous *because*-paraphrase, as the following examples illustrate:

- (72) John did not win the race. He would have celebrated
(73) ?Because he would have celebrated, John did not win the race
(74) The stadium was not built. It would have had 40 000 seats
(75) Because it would have had 40 000 seats, the stadium was not built.

Although the latter example is felicitous, it represents only under a special reading a synonym of the discourse. This reading can for instance be based on the accommodation of an underlying law according to which it is forbidden to build stadiums with 40 000 seats.

CHAPTER 1. THE DATA

it should not be, since the discourse initial sentence is not negated, and thus the counterfactual consequent should not find a salient counterfactual antecedent in its context to complete a conditional.

The first point to be observed in connection with the *because*-paraphrase of this set of data is again the parallel in felicity. Whenever the original mini-discourse is acceptable, also the corresponding *because*-sentence is, whereas in case the example is infelicitous, also the *because*-sentence is not. (For accessibility the examples are repeated again here.) The first example follows the NinA-pattern: the main clause (corresponding to the initial sentence of the discourse) is negated, and thus provides with a NinA that could act as an implicit antecedent to the counterfactual consequent in the subordinate (second) clause:

(76) John did not have surgery. He would have died.

(77) Because he would have died, John did not have surgery.

The structural pattern of the second example challenges the NinA-account: the main clause contains a sentential negation, and thus provides with a NinA, but still the sentence is not felicitous:

(78) #John did not have surgery. He would not have died.

(79) #Because he would not have died, John did not have surgery.²³

The third example of a *because*-paraphrase challenges the NinA-account once more from a different direction. While the main clause is not negated, and thus does not provide a NinA to complete the counterfactual, the *because*-sentence is nonetheless felicitous, although the NinA-account predicts its infelicity.

(80) John had surgery. He would have died.

(81) Because he would have died, John had surgery.

Contrary to this in the fourth case (that differs from the third only in that its subor-

²³Also this paraphrase and the one given in (83) allow for a second reading to the effect that it is understood that in order to guarantee the patient's death surgery was (resp. was not) carried out.

CHAPTER 1. THE DATA

dinate clause is negated, while in (81) it is not negated) the NinA-account seems to be confirmed. The main clause is again positive, and thus can not carry a NinA. And, as initially predicted, the *because*-sentence as a whole is not acceptable.

(82) #John had surgery. He would not have died.

(83) # Because he would not have died, John had surgery.

1.7 Exhaustivity as a formal constraint on discourse felicity?

Returning to the mini-discourses of the original point of departure we will see that a reconsideration of the data delivers a new perspective that will eventually allow to formulate a solution to the puzzle. The starting point is provided by the question what conditions are imaginable that could have brought about the event hypothesized in the consequent of the mini-discourse below?:

Positive consequent

(84) John did not have surgery. If he had had surgery, he would have died.

Conceived somewhat more abstractly, the underlying question is what characterizes the set of possible antecedents of the counterfactual conditional²⁴:

If x had been the case, John would have died.

The answer is of course that there is a multitude of possible scenarios that might bring about the effect described in the consequent: John could have and would have possibly died not only if he had had surgery, but also if he had had a heart attack, *or* if he had had a car accident, *or* if he had caught the Ebola virus, *or* if he had been murdered by an assassin, *or* if . . . and so on. We see that only one of the possible conditions needs to hold in order to effectuate the consequent and to

²⁴The question that is employed here is probably best conceived as the *question under discussion*, abbreviated as *QUD* (cf. Roberts (1996b); but see also van Kupperfelt (1996) or Ginzburg (2012) for similar concepts.)

CHAPTER 1. THE DATA

render it true. Thus, the set of possible antecedents is *disjunctive* (and within the counterfactual realm essentially unrestricted).

Negative consequent

Now let us consider next, whether this holds also for the case of a negative consequent:

If x had been the case, John would not have died.

Here, again, it is easy to imagine a plethora of possibilities that present possible antecedents to the consequent. The difference to the non-negated case however is that for John not to die not only he must have had surgery, but also he must not have had an accident, *and* he must not have been shot at, *and* he must not have caught the Ebola virus, and he must not have had . . . and so on. This means, that in this case negation in the consequent demands to exclude *all* possible eventualities from consideration. The logical structure of the premise set is thus that of a *conjunction*. In its constraining force on the premise set the negated consequent behaves opposite to the case of the positive consequent. While the latter determines an existential constraint on the domain of possible antecedents, the former exerts a universal demand on the set of premises.²⁵

A tentative assessment of the observation

When assessing what to make out of this finding with respect to the (in)felicity of the data, one might ponder the following argument²⁶:

²⁵Recall that the shift of attention away from the (availability of an) implicit antecedent, resp. its source of origin in the initial context, was motivated by the observation that for the felicity of the data under consideration it apparently did not matter whether the discourse initial sentence was negative or not. See again the table given on page 20 which summarized the data and its behavior. Hence, it suffices to ponder just upon the polarity of the consequent, and to disregard the polarity of the initial sentence.

²⁶We will see that ultimately the argument has to be revised. Felicitous examples like

(85) John did not drink any alcohol. He would not have passed the impending drug test.

with a negative initial sentence, or

CHAPTER 1. THE DATA

The discourses under discussion are felicitous, whenever the counterfactual consequent can link up to a contextually given, implicit antecedent. The antecedent is selected from a set of possible alternatives: *John would have died if he had had surgery, or if he had had a heart attack, or if . . .*. An important feature of this set of possible premises is that it is essentially unrestricted. The context is never so definite that it could determine the exact members of the premise set. After all, the linguistic agent is ultimately free in what she assumes²⁷: besides the more common premises just mentioned one might as well assume more exotic conditions under which *John would have died*, like for instance *if he had caught the Ebola-virus* or *if he had been struck by lightning*. While these conditions may not be very probable, they are still possible. And this is everything that is needed for making an assumption.²⁸

With the introduction of negation in the consequent neither the context, nor the possible premises themselves are changed. What changes is rather the logical structure of the premise set. It is no longer *disjunctive*, but becomes, as already observed, *conjunctive*: for John not to die he (i) must not just not had surgery, but also he (ii) must not have had a heart attack, and he (iii) must not have been struck by lightning, and . . . This means that in the case of a conjunctive premise set not just one instantiation of the possible antecedents needs to obtain, but *all* of the possible premises are to be excluded. However, considering that premises

(86) The pilot took the compass with him. He would not have found his way home.

with an affirmative initial sentence apparently contradict the hypothesis that negation somehow lexicalizes the exhaustivity constraint put forth on the set of possible antecedents. Also a felicitous discourse like

(87) John won the race. He would not have celebrated.

poses a problem for the suggested account. However, the core idea deriving from the observation – namely the distinction between disjunctive and conjunctive premise sets (which also marks the difference between sufficient and necessary conditions) – will provide with the key to the account of the data under consideration.

²⁷I take this freedom of assumption not to be total. What could possibly be assumed is restricted by the conceptual scheme entertained.

²⁸This does not just hold for counterfactual assumptions, but also for the assumption expressed in the antecedent of indicative conditionals. Thus, the matter is not one of mood, as one might think.

CHAPTER 1. THE DATA

are free to be assumed this seems to be a demand that can not be met. Indeed, the semantics demands to exclude what the pragmatics determines to be in principle not excludable: possibility. With the introduction of negation the semantics and the pragmatics of the counterfactual consequent thus clash in a way that can not be resolved. This eventually leads to the infelicity of the data.²⁹

1.7.1 Wrapping things up

Yet, the analysis still needs to be driven further. A rather important question immediately arises from this proposal. One needs to account for the difference in felicity between discourses that contain an implicit antecedent versus those in which the antecedent is explicitly given.

How could it be explained that the infelicitous discourse repeated here in (88) becomes felicitous as soon as the antecedent is made explicit (as in (89)), although both arguably specify the exact same counterfactual premise *If John had had surgery ...* (which in case of (88) is given implicitly as a NinA).

(88) # John did not have surgery. He would not have died.

(89) John did not have surgery. If he had had surgery, he would not have died.

Apparently, in the passage from an implicit to an explicit antecedent in the discourse the counterfactual domain of the antecedent is restricted and the exhaustivity constraint met (which in turn results in the felicity of the discourse). So, the question is how is this restriction achieved, and what grammatical object is responsible for it?

With respect to the latter part of the question I will rely on the obvious: what constitutes the difference between the antecedent being implicit vs. it being explicit is that the latter is actually uttered, while the former is not. Crucial for the

²⁹One might drive the point even further. When the conclusion that *John would not have died* rests on the exclusion of any conditional premise, the matter becomes literally unconditional. But to express an unconditional claim in a conditional statement seems absurd and a contradiction in itself.

CHAPTER 1. THE DATA

felicity of the non-elliptic discourse is that – although the expressed antecedent remains but just one of a multitude of possible alternative premises – these alternative antecedents are not mentioned in the discourse. But whenever none of these possibilities is deemed worth the effort of mentioning, one might reason that these alternatives to the actually uttered antecedent must be false and thus irrelevant. (Which in turn satisfies the exhaustivity constraint.) Linguistically this difference in status (and, consequently, in felicity) can be captured by assuming an illocutionary operator ‘UTTER/ASSUME’ that is an adjustment of Krifka’s ‘ASSERT’-operator (as formulated for instance in Krifka (1999), p. 10) to the peculiarities of conditional sentences.³⁰

UTTER/ASSUME(M, A, c) (a sentence with meaning M and alternatives A in a context c is uttered/assumed):

- the speaker claims M (in c)
- for every alternative $M' \in A$, $M' \neq M$, the speaker does not claim M' (in c).

Assuming that the antecedent of a counterfactual conditional sentence always derives from a set of competing alternative premises (which is reciprocal to the assumption that *iff* is not a lexicalized expression) this allows to motivate the choice of a particular antecedent, since, as Krifka argues, the following holds:

There must be a pragmatic reason why the speaker introduces alternative propositions without actually claiming them, reasons that are obvious to the hearer. There are, essentially, two types of reasons for the cases that are relevant here:

- It may be that M' is more informative than M (in c). The reason in this case is obviously that the speaker lacks evidence for M' , and perhaps even has evidence that M' is false. That is, uttering M' instead would violate the maxim of Quality.

³⁰An alternative approach to the issue (adopted from von Stechow (2001)) will be discussed later on in section 4.2.4 on page 134 ff., where a related issue is examined.

CHAPTER 1. THE DATA

- It may be that M' is less informative than M (in c). The reason in this case is obviously that the speaker prefers M because it gives more information, in c. That is, uttering M' instead would violate the Maxim of Quantity.
(Krifka (1999), p. 10)

How this applies to the present concern becomes obvious when recalling that for the counterfactual consequent *John would not have died* the logical structure of the set of possible antecedents is that of a conjunction. For John not to die he must not just have had surgery, but he must not have had an accident, and he must not have been shot at etc. The possible alternatives thus conjunctively add up. This however means that for every alternative M' to M, M' is immediately *more informative* than M alone: assuming that M is *John had had surgery*, the 'first' alternative to it in order to guarantee that he would not have died is that *John had had surgery and had not had an accident* to which an even 'better' alternative is that *John had had surgery and had not had an accident and had not been shot at* etc. Hence, whenever these alternatives to the information that *John had had surgery* are not mentioned, Gricean reasoning allows to conclude, as Krifka points out in the first bullet point of his elaboration above, that these may be assumed to be false. This in turn however sets a limit to exhaustivity, because if the uttered proposition would not exhaust the set of relevant alternatives the Maxim of Quantity would be violated.

On this account what is actually uttered and what is merely induced have different restrictive powers due to the presence, respectively absence of an UTTER/ASSUME-operator. In the absence of the UTTER/ASSUME-operator the domain of the antecedent can not be restricted appropriately, and the conditional is not strengthened.³¹

³¹Note that the point also extends to the particle *otherwise*. As a discourse mirror it serves multiple functions. It not only provides with the necessary semantic content, but acts as the link for the illocutionary operator. That expressedness has besides the discussed pragmatic effects with respect to felicity also important semantic repercussions (in form of an influence on semantic acceptability) will become clear later on, when the formal semantic analysis is approached.

1.8 Summary

The main task of the chapter was to introduce the data that present the empirical basis of this thesis and to formulate the observations and problems that derive from the examples. The starting point is provided by observations made in [Veltman \(2005\)](#) that yield what I call *Veltman's puzzle*. This puzzle connects to the discourse behavior and felicity conditions of certain kinds of counterfactual discourse. These discourses consist of two sentences: some (atomic or negated) factual proposition φ expressed by a discourse initial sentence S_1 that is followed by a counterfactual conditional S_2 the antecedent of which opposes the preceding factual information. (Actually, in the discourses under discussion the discourse initial sentence corresponds to the presupposition of the counterfactual's antecedent.) The general form of the sequence is thus: $[\varphi]_{S_1} . [\neg\varphi \Box\rightarrow \psi]_{S_2}$ with the boxed arrow ' $\Box\rightarrow$ ' formalizing that the conditional is a counterfactual conditional.

What is peculiar to this sequence is a difference in discourse felicity once the antecedent of the counterfactual is left implicit. Whenever an acceptable sequence with implicit antecedent (given below by ...) is inverted, it becomes infelicitous. Or, more formally exemplified at hand of an instantiation, while $[\neg p] . [\dots \Box\rightarrow q]$ is felicitous, its opposite is not: $\# [p] . [\dots \Box\rightarrow \neg q]$. Thus, the initial problem is a pragmatic problem. One has to account for the felicity conditions of implicit counterfactuals. This necessitates an analysis of their interplay with context and negation.

That negation plays a special role for the felicity of the data is recognized in a first, tentative account. Key to this is the formulation of the concept of *negation-induced alternatives* (NinAs). The gist lies in the hypothesis that a negated discourse initial sentence $\neg S_1$ induces as a counterfactual alternative to the expressed sentence $\neg p$ the proposition p into the context. This proposition in turn complements as an implicit antecedent the explicitly given counterfactual consequent. Thus the conditional is completed and becomes interpretable. In case S_1 is not negated no such counterfactual alternative is provided to which the counterfactual

CHAPTER 1. THE DATA

consequent could link up to. The conditional thus remains incomplete and hence not interpretable. Which would explain the infelicity of the discourse, as attested by the first set of data.

That the concept of the negation-induced alternatives has empirical substance is confirmed at hand of data involving the particle *otherwise* that already Veltman mentions. Based on this a generalization and extension of the empirical concern is attempted, and other alternative inducing linguistic phenomena (focus and alternative questions) are briefly discussed as possible contexts for counterfactual consequents. Although a first set of data seems to support the envisaged account, with the consideration of more examples the sketch becomes increasingly problematic. Indeed, when returning to the original type of data that involve sequences of fact-counterfactual (conditional) it becomes clear that the account in terms of the NinAs is ultimately not tenable. The account is contradicted twice. First, by examples that produce a NinA, but that still turn out to be infelicitous, and next by examples that do not deliver a NinA, but that nonetheless are felicitous.

However, a shift in attention towards the peculiar discourse relation that characterizes the majority of the examples provides with a promising lead. Prevalent to the data is the understanding that the discourses express some kind of (loosely understood) ‘causal’ nexus which is attested by the possibility of a paraphrase (of most) of the examples in terms of a *because*-sentence. The question what conditions have to obtain for the consequent to turn out true yields a significant insight. The logical structure of the possible premise sets from which the implicit antecedent derives is either *conjunctive* or *disjunctive*. In case of a conjunctive set all possible ‘causes’ to the consequent would have to be exhausted in the evaluation of the consequent, while in case of a disjunctive set a single witness suffices to determine whether the consequent is supported or not. Since – at least for the discussed particular example – the logical character of the premise set switches with the occurrence of negation in the consequent from being disjunctive to being conjunctive, it is hypothesized that negation (when occurring in the consequent) exerts an exhaustivity constraint on the set of possible antecedents. The main idea regarding the infelicity of the data is that at this point semantics and pragmatics

CHAPTER 1. THE DATA

clash and can not be reconciled. The semantic constraint put forth by the counterfactual consequent, namely to exhaust in its evaluation the (implicit) set of all possible antecedents, overstrains a fundamental pragmatic principle connected to model construction in the counterfactual domain: anything is possible, nothing can be excluded.

Chapter 2

The Empirical Vicinity

Outline of the chapter

The main part of the chapter is devoted to giving an overview on the concept of modal subordination and the motivating data. In sections 2.1.1 - 2.1.2 certain problems regarding anaphora in two-sentence mini-discourses are laid out that connect to the presence, resp. absence of (intensional) quantifying elements in the second sentence. These issues resolve, once it is recognized in section 2.1.3 that the discourse-initial first sentence provides with an implicit antecedent to the second sentence in form of a contextually determined restriction. In the following two sections the account is extended to data that include intensional predicates (2.1.4) and presupposition triggers (2.1.6). This leads to a more general formulation with respect to what structures may give rise to modal subordination in section 2.1.7. After this the concern shifts in the conclusion of the discussion of modal subordination to more material issues: in 2.2 and its subsections 2.2.1 and 2.2.2 the problem how exactly the context and the involved intensional operator interact is addressed, the question being whether the link is semantic (in terms of entailments), or pragmatic (in terms of accommodation). The second half of the presentation of the empirical vicinity of the Veltman's puzzle (section 2.3) consists of a brief discussion of the data and the account presented in Schueler (2008), who studies a type of implicit counterfactual conditionals that appears to be complementary to the data I am concerned with.

2.1 Modal subordination

The link between the data discussed so far and the following presentation of what has been deemed *modal subordination* since Roberts (1987) is provided by the observation that in both cases implicit, contextual information plays a decisive role in the constitution of felicitous discourse. Indeed, it turns out that for both, the discourses which give rise to Veltman's puzzle and the classical modal subordination data, the recourse to readily identifiable parts of the implicit context guarantees not only discourse coherence – i.e. model consistency –, but also, and arguably more fundamentally, discourse felicity (in terms of interpretability). While the phenomena seem to be closely related, the discussion will reveal that the data as discussed so far add a new, interesting facet to the modal subordination picture. It allows to identify as of yet unrecognized constraints on the interaction of implicit and explicit discourse and the constitution of larger linguistic units at the intersection of semantics and pragmatics.

The account of modal subordination takes an at first sight rather surprising point of departure that seems quite unrelated to the phenomena encountered so far. It is problems concerning anaphora and its relation to discourse felicity. However, the relevance for the actual concern – Veltman's puzzle – will become clear soon. The presentation to follow is based on Roberts (1996a).

2.1.1 Preparing the ground: anaphora and discourse felicity

The problems that Roberts takes as her point of departure in Roberts (1996a) are characterized by an undeniable tension between theoretical account(s) and empirical reality. On the theoretical side we have various analyses of anaphora that help to identify and formulate rather 'robust' constraints on the subject (cf. Roberts (1996a), p. 216). On the other side however, there are data that clearly challenge these otherwise well supported and established generalizations on anaphora. The dilemma is that these data are felicitous where theory would predict them not to be. So, the question of course is what to do with this? Ignore the data? Or revise the otherwise rather sound and proven theories? It turns out that a third option ex-

CHAPTER 2. THE EMPIRICAL VICINITY

ists that is theoretically conservative insofar as it maintains the established results, but that is at the same time empirically accurate by giving a motivated analysis of what is really going on and how to account for it. The lesson to be learned from Roberts' discussion of the topic is to not always take things for what they seem to be. The starting point however is rather straightforward. It is provided by the following data given in Roberts (1996a), p. 216:

- (1) # **Every** frog that saw an insect ate it. It was a fly.¹
- (2) # You **should** buy a lottery ticket and put it in a safe place. It's worth a million dollars.

While intuitively the discourses are clearly infelicitous, the source of this remains unclear at first sight. Although it suggests itself to locate the issue with the anaphorical relations expressed in the discourses, these relations seem at the same time unquestionable: the underlined indefinite NP acts as the antecedent to both occurrences of the pronoun *it* in the data (which are again underlined). So, what is going wrong here? The answer is provided by the following “descriptive generalizations” on anaphoric accessibility (Roberts (1996a), p. 216):²

¹The marking follows Roberts' example. **Bold** font identifies quantificational expressions, while intuitive reference relations are highlighted by underlining of the respective constituents.

²An important qualification made by Roberts regarding the generality of the constraints needs to be mentioned (cf. Roberts (1996a), p. 244, endnote 2): definite and indefinite noun phrases defy in their ‘anaphoric potential’, as Roberts puts it, the given description: “Unlike NPs with determiners like *every* or *no*, definite and indefinite NPs may be anaphorically accessible across discourse.” While Roberts does not bother to give an illustration of the issue, for indefinites one may refer to Groenendijk and Stokhof's classic example (cf. Groenendijk and Stokhof (1991), p. 41 ff.):

- (3) A man walks in the park. He whistles.

Although intuitively the pronoun *he* in the second sentence is anaphorically linked to the NP *a man* in the first sentence, in a translation to classical predicate logic that strictly observes the *sentential scope constraint* the pronoun lies outside the scope of the indefinite's translation as an existential quantifier. –Which contradicts the *scope constraint on anaphoric relations* (at least as long as one considers the indefinite not to be a constituent of the NP itself). For definite descriptions a similar problem arises. Substituting into the example the definite noun phrase *the president of the United States* for the indefinite *a man* generates a felicitous sequence of sentences in which the pronoun in the second sentence has access to its antecedent across a sentence boundary, and hence also violates the *scope constraint on anaphoric relations*. Thus, the notion of a *quantificational element* that plays such a pivotal role in the formulation of the constraints has to be understood

CHAPTER 2. THE EMPIRICAL VICINITY

Scope constraint on anaphoric relations: If NP x is anaphorically accessible to NP y , then any quantificational elements which have scope over x have scope over y as well.

Sentential scope constraint: The maximal scope of a quantificational element is the sentence in which it occurs.

When reconsidering the examples in (1) and (2) with respect to these principles it becomes clear in how far the examples violate the given constraints. Assuming that a sentence boundary is determined by the potential of an expression to bear a truth value, it can indeed be observed in the data that the occurrence of the pronoun *it* in the second sentence lies outside of the scope not just of the indefinite NP in the first sentence, but also outside of the scope of those quantificational elements that further occur in the first sentences: *every* and *should*. This sets this occurrence of *it* apart from its first instantiation in the initial sentence that lies within the scope of the quantifiers present.

On this ground one may thus argue that although intuitively the pronoun *it* in the second sentence refers to the indefinite NP of the first sentence, it is nonetheless formally not bound, and thus its interpretation not resolvable. In turn the discourses are infelicitous, because they are incoherent. Or, in Roberts' words: "The NP is anaphorically accessible only to the first of two pronouns, the one which is also under the scope of the quantificational element, and not to the second pronoun, which is outside that scope under the assumption that quantifier scope is sentence-bounded. Because the second sentence in each discourse contains a pronoun with no potential antecedent, the sentence is uninterpretable and hence infelicitous, as indicated by the "#" sign." (Roberts (1996a), p. 216)

with certain reservations and arguably does not include definites and indefinites. More problematic to Robert's account are data like

(4) Every frog that saw an insect ate it. It tasted deliciously.

This issue shall however not be pursued any further.

2.1.2 More data and a problem for the established constraints

The presentation so far gave in a nutshell the theoretical and empirical background into which the following data ought ideally to be integrated. The new examples (cf. Roberts (1996a), p. 216) deviate from those already introduced in only a detail: the second sentence now contains a quantificational element. However, this amendment has drastic effects regarding acceptability, and thus has repercussions for the account as a whole. Unlike in (1) and (2) (repeated here in (5) and (7) to highlight the contrast) the data in (6) and (8) represent coherent and felicitous discourses. –Although in terms of those structural aspects alluded to in the accessibility-constraints the discourses in fact have not changed. In both cases the initial pronoun in the second sentence lies outside the sentence boundary and quantifier scope given with the first sentence.

- (5) # Every frog that saw an insect ate it. It was a fly.
- (6) Every frog that saw an insect ate it. It disappeared forever.
- (7) # You should buy a lottery ticket and put it in a safe place. It’s worth a million dollars.
- (8) You should buy a lottery ticket and put it in a safe place. It might be worth a million dollars (if you were lucky).

Based on this observation the conclusion is that apparently the insertion of a quantifier like *forever* or a modal like *might* (which, sticking to standard semantic analysis, will be treated as quantificational just as well) facilitates anaphoric accessibility across sentences, i.e. cross sentential anaphora. This however runs of course against the insights underlying the given constraints on anaphoric accessibility –which is rather disturbing for those theories of anaphora that incorporate the constraints. While they can explain the infelicity of the original data, they cannot explain the felicity of the quantified resp. modalized variants just given. Indeed, the prediction is that these examples ought to be infelicitous, too, since the anaphoric pronoun in the second sentence of the data does not lie within the scope of the indefinite nominal contained in the first sentence. Or, as Roberts puts it: “These discourses and many others like them are perfectly acceptable to most

CHAPTER 2. THE EMPIRICAL VICINITY

speakers, despite the apparent anaphoric relation between a pronoun in the second sentence and an indefinite NP which is under the scope of the [...] quantificational element in the first.” (Roberts (1996a), p. 217)³

2.1.3 The key to what is going on in the data and its relation to Veltman’s puzzle

The key to what is happening here is found when contemplating upon what is actually expressed in the data. While on the surface the data may have only changed in terms of the presence of a modal or other quantifier, their reading however is intuitively richer than this. Indeed, the data contain a good amount of implicit information that is – just as with the discourses of Veltman’s puzzle – provided by the context, i.e. by the initial sentence. And it is, again similar to the puzzle-discourses, to this contextually given elements to which the modal or intensional quantifier eventually links up to. A paraphrase in which the otherwise implicit information is spelled out makes this clear. Example (9) is the non-elliptic variant of example (6), while (10) is a fully explicit rendering of (8).

- (9) Every frog that saw an insect ate it. The insect that was seen and eaten by each frog disappeared forever.⁴
- (10) You should buy a lottery ticket and put it in a safe place. The lottery ticket you should buy and which you should put in a safe place might be worth a million dollars (if you were lucky).

A rough and schematic translation that captures the structural difference between e.g. examples (1) and (6) will make the impact that the introduction of an (intensional) operator resp. modal has on the felicity of the data perspicuous. Although also in (1) it is intuitively clear that the pronoun in the second sentence refers to the indefinite NP in the first sentence (i.e. to contextually given information), the semantic structure of the second sentence does not support this link. The second sentence that follows the first sentence translates in case of example (1) to an open

³The quantificational elements in question are *every* and *should*.

⁴I am solely sticking to a distributive reading here.

CHAPTER 2. THE EMPIRICAL VICINITY

formula:

[sentence 1] $[Fly(x)]_{=sentence2}$

Contrary to this in case of (6) the second sentence is a closed formula, since the semantics of the adverb *forever* changes the logical structure of the second sentence. The decisive point however is not just that *forever* as an intensional operator closes the formula, but that it furthermore introduces a slot for a restrictor that fixes the domain of interpretation for the information given in its scope via the verb *disappear*.

[sentence 1] $[FOREVER (..._{restrictor} \rightarrow DISAPPEAR(x))]_{=sentence2}$

What is crucial to be noted is that within the context of the first sentence a proper domain restriction (in form of a complex NP) for the quantificational element in the second sentence becomes available that is otherwise lacking. Indeed, while no explicit restrictor is given in sentence 2, the context implicitly delivers the information about the restriction on the domain. This in turn renders the discourse felicitous. Thus, in sum it is the different structure and logical form that sets the felicitous examples apart from the infelicitous examples in (1) and (2).⁵

On basis of this observation one can then give a straightforward account of the data in terms of the established constraints on anaphora: the restriction given by the context figures as the antecedent to the potentially anaphoric pronoun in the second sentence. Since both, the antecedent NP, and the anaphoric pronoun are located within the scope of a quantifier within the boundaries of a single sentence, the accessibility constraints are not violated. This means that what appears at first sight to be cross sentential anaphora happens to be under this analysis a common case of anaphora within a single sentence. (With the antecedent of course deriving

⁵Just note that in case of example (10) the restriction of the might operator is arguably given only by the first proposition of the first sentence. The second conjunct acts under this reading as a rhetorical element, as the following paraphrases show: *Because it might be worth a million dollars you should buy a lottery ticket. And hence you should (also) put it in a safe place* versus ??*Because it might be worth a million dollars, you should put a lottery ticket in a safe place. And hence you should (also) buy it.* This issue is however not addressed by Roberts.

CHAPTER 2. THE EMPIRICAL VICINITY

from the context.)

The problem is thus not so much how to account for the anaphoricity in the data, but how to account for the exact relation between the context and the restriction of the domain of interpretation in the subsequent sentence. This question will be addressed in section 2.2. However, before doing so the empirical base shall be extended, and other data that apparently show similar behavior shall first be introduced.

2.1.4 Extending the empirical base, part 1: intensional contexts

In the coming discussion I will continue to follow Roberts (1996a), who argues with reference to the already introduced data that there is a “closely related phenomenon involving propositional attitude verbs and ‘world-creating’ predicates more generally.” (Roberts (1996a), p. 217) To support her point she presents (with reference to Baker (1966), Lakoff (1970), Jackendoff (1972), Montague (1973), Karttunen (1974), McCawley (1981), and Heim (1992) (cf. (Roberts (1996a), p. 217)) the following examples (in which the “world-creating” predicates are printed in **bold**, while the anaphoric elements are underlined):

- (11) Jan **expected** to get a new puppy soon. She **intended** to keep it in her backyard.
- (12) John **wants** to catch a fish. He **plans** to eat it for supper.
- (13) Alice **fears** there’s a squirrel in her kitchen cabinets. She **hopes** to trap it alive and turn it loose outside.

As before the apparent anaphorical relation between the pronoun in the second sentence and its antecedent in the first sentence clearly violates the constraints on anaphoric accessibility: pronoun and antecedent do neither occur within the scope of the same operator, nor do they appear within the same sentence. Moreover, both are located within the scope of different opacifying verbs, like *expect* and *intend* in the first case. Nonetheless, the discourse is unproblematic and the

CHAPTER 2. THE EMPIRICAL VICINITY

anaphorical relation established. Again, it seems that this is a case of cross sentential anaphora that contradicts acknowledged theoretical insight. Yet, again, the issue resolves: closer contemplation upon the data reveals a striking parallel to the examples discussed before. Also in the present cases the second sentence comes with an implicit contextual restriction. This restriction is given by the preceding sentence and ultimately provides the anaphoric antecedent to which the pronoun contained in the second sentential projection actually refers to. For the first example the worlds that are subject to the agent's intentions are just those worlds in which her expectations turned to fact and she actually has already obtained a new puppy. Only for those worlds in which Jan is in the possession of a new puppy did she intend to keep it in the backyard. Anything else would intuitively be untenable. Similar for the other discourses. In the second example the worlds that the subject considers in his plans are solely those worlds in which the premise is realized and he in fact managed to catch a fish, while the second sentence of the third discourse is concerned only with those worlds in which the agent's fears have actually come true and there is indeed a squirrel in her kitchen. Only in this scenario are her hopes of trapping the squirrel and letting it loose founded at all. Indeed, in both cases the otherwise implicit restriction can be spelled out without altering the meaning of the discourses as a whole:

- (14) John wants to catch a fish. He plans to eat the fish he caught – if he actually catches one – for supper.
- (15) Alice fears there's a squirrel in her kitchen cabinets. She hopes to trap the squirrel in her kitchen cabinets – if there indeed is one – alive and turn it loose outside.

Roberts eventually relates the phenomenon to a single source: “All the felicitous examples except (1') [appearing here as example (6), D. B.] involve either a modal, an opacifying predicate, or an adverb of quantification in the second sentence.”⁶ (Roberts (1996a), p. 218 - 219) Taking this perspective one may indeed

⁶Her example (1'), which is given here as (6) is arguably not so much an exception as Roberts takes it to be. (In support of her point of view she refers on p. 219 to Roberts (1987), and Roberts (1989).) *Forever* may be considered intensional if one assumes that it quantifies over all possible temporal successors to the actual world.

CHAPTER 2. THE EMPIRICAL VICINITY

subsume the data also under the heading of ‘anaphora in intensional contexts’, as Roberts puts it in the title of [Roberts \(1996a\)](#).

2.1.5 Preliminary summary

Bringing the phenomena introduced so far into perspective, the main points of the discussion are, first, the fact that the insertion of a quantificational element or of an intensional predicate in the second sentence provides with a link to the preceding context, and, second, – on the assumption of the empirical correctness of the constraints on anaphoric accessibility – the observation that in turn (parts of) the given context – just as with the discourses of Veltman’s puzzle – actually enter the logical form of subsequent discourse and, although remaining implicit, play a decisive role there in facilitating discourse felicity. In fact, at least in case of the first set of examples, the context as given by the initial sentence is of twofold importance: it not only delivers a restriction to the operator in the second sentence, but this restriction in turn acts as the antecedent to whatever pronouns in the second sentence are dependent on an anaphorical interpretation.⁷

2.1.6 Extending the empirical base, part 2: modal subordination and presupposition

Following Roberts the same mechanics arguably underlies also another phenomenon to which she draws a final connection. It is presupposition satisfaction in discourse. A brief presentation of this topic will complete the empirical overview. Considering the examples below the observation is that – just as before – the

⁷The data presented by Roberts seem to suggest that the presence of an intensional predicate in the initial context is indispensable. This is not the case, as the following amendment exemplifies. Dropping the first intensional predicate does not render the discourse infelicitous:

(16) John buys a fish. He plans to eat it for dinner.

But see again my footnote 2 on page 36 for the influence of indefinite (and definite) articles on anaphoric accessibility. On the other side, as to be expected, dropping the intensional predicate in subsequent discourse yields an infelicitous discourse:

(17) #John wants to catch a fish. He (will) eat(s) it for dinner.

CHAPTER 2. THE EMPIRICAL VICINITY

second sentences – which contain presupposition triggers – are interpretable or felicitous only within the context of the initial discourse.

While successful anaphora resolution runs parallel to the next examples in that it presupposes the accessibility of an appropriate antecedent, Roberts argues that in the data to follow (cf. Roberts (1996a), p. 219, bold font and underlining as in the original) “the second sentence carries a presupposition which, in the discourses in question, is only satisfied by material which is under the scope of the quantificational element or predicate in boldface in the first sentence”. (Roberts (1996a), p. 219) Indeed, the gist of her argumentation (and the motivation for the term ‘modal subordination’, as introduced in Roberts (1987)) is that these presuppositions are “semantically subordinate to irrealis propositions considered in previous discourse.” (Roberts (1996a), p. 219) –Which means that the presuppositions are in this case entailed by the preceding (counterfactual) discourse. The (focus induced) presupposition for (18) below is that there is something else besides coffee that Fred buys. For (19) and (20) the presupposition is given by the truth of the complement of the factive verbs *discover* and *regret*, which is for (19) that Maxine is indeed capable of building things, and in case of (20) that Mary is attending grad school.⁸

- (18) **Usually** Fred buys a muffin every morning and eats it at the office. He buys [a cup of coffee]_F, too.
- (19) Maxine **should** become a carpenter. Her friends **would** discover she could build things, and she’d be very popular on weekends.
- (20) Mary is **considering** getting her Ph.D. in linguistics. She **wouldn’t** regret attending graduate school.

⁸Note that these presuppositions come with a very important distinction: while the discovery that Mary could build things may include already the factual world in the sense that one may assume that having the capacity to build is not necessarily dependent on being a trained carpenter, no such *global* validity connects to the second presupposition. In this case the presupposition is linked exclusively to the irrealis and excludes the factual world: Mary is not yet attending grad school. Only in case she were to pursue a Ph.D. in linguistics – which she is not yet doing – were she to attend grad school. This is a case of a *local* accommodation of a presupposition, in the sense that the presupposition connects only to some hypothetical scenario that excludes the factual world.

CHAPTER 2. THE EMPIRICAL VICINITY

What Roberts alludes to in the quote above becomes clearer once it is pointed out that just as in the case of the anaphorical data, the second sentence in the examples becomes interpretable only within the context of the given initial sentences. Without these it will not become clear what presupposition the particle *too* in (18) rests on, nor what motivation there is for what is expressed in the second sentences of (19) and (20). For all these cases a lack of the necessary information results (my interpretation) in incomprehensibility of what is expressed in the second sentence.⁹

But note that arguably this incomprehensibility is due to the absence of a condition that could describe a motivational nexus or context for what is expressed in the sentences containing *would*. One may object rather that the *would* sentences are consequents without an antecedent. If one follows this argument the whole point that Roberts tries to make becomes questionable. Especially her remark that “[t]hese presuppositional parallels with the earlier anaphoric examples support the analysis of anaphora as essentially presuppositional, as argued e.g. by Heim (1982). [Footnote omitted, D.B.] Hence, a more appropriate title for this paper might be “Presupposition satisfaction in intensional contexts”. But for simplicity I will focus on the examples involving anaphora.” (Roberts (1996a), p. 220) –What these presuppositional parallels actually are is regrettably not explicated in more detail. More important than this however is at this moment the summary

⁹Note that at least in case of (19) the second sentence at the same time ‘motivates’ the deontic necessity expressed in the initial sentence, as Roberts points out. She elaborates on the data: “Following Heim (1992), in (12) [given here as (18), D.B.] *too* presupposes that it is known by the interlocutors that there is some salient entity which could replace the property denoted by the focused constituent *a cup of coffee* while preserving truth, i.e. there is something else which the interlocutors already know that Fred buys. Further, the bare present tense with a non-stative verb is interpreted as the generic mood, in recent literature often formalized with an adverbial operator (see Carlson (1977) and Wilkinson (1991)). In (13) [given here as (19), D.B.] *discover* presupposes the truth of its complement. But, under the assumption that Maxine isn’t already skilled as a carpenter, this presupposition is only satisfied in those deontically ideal worlds where she becomes one, verifying the proposition introduced under the scope of *should* in the previous sentence. The modal auxiliary *would* will have to range over such worlds in which that proposition, that Maxine is a carpenter, is true, in order to satisfy the factive presupposition of *discover*. *Regret* is similarly factive; in (14) [here: (20), D.B.] it is under the scope of a modal, but (assuming Mary hasn’t yet gone to graduate school) its presupposition is only satisfied in those worlds which realize what she is considering, introduced in the previous sentence.” (Roberts (1996a), p. 219 - 220)

CHAPTER 2. THE EMPIRICAL VICINITY

that Roberts gives herself of what she regards to be the common denominator in the data. This general scheme will be presented next.

2.1.7 Recapitulating and connecting the observations: the general scheme

Roberts' observations are summarized in the following logical forms that she argues to underlie the data. They follow a single scheme, as she remarks:

The schema in (15) [(21) here, D.B.] captures the general logical form of examples displaying modal subordination, while that in (16) [my (22), D.B.] gives the anaphoric sub-type:

- (21) operator_S ... π ...
operator_{intensional}_S ... presupposition of π ...]
- (22) operator_S ... NP_{x,[-def]} ...]
operator_{intensional}_S ... NP_{x,[+def]} ...]

In each discourse there are two sentences, each containing a wide scope operator. In the second sentence, the operator is intensional – the operator logically associated with a propositional attitude predicate, a modal auxiliary, or an adverb of quantification, ranging over situations – though this operator is sometimes implicit, as in (1') [given here as (6), D.B.] and (12) [= (18), D.B.] In the general case, some element π under the scope of the operator in the first sentence is presupposed by an element which has narrow scope under the intensional operator in the second sentence. (Roberts (1996a), p. 220)

With regards to the initial anaphoric data in which an indefinite NP is anaphorically accessible to a (definite) pronoun Roberts continues to elaborate: “In the sub-case of interest, the presupposition is one of the familiarity of some entity introduced by an indefinite NP_x in the first sentence; i.e. there is an anaphoric relationship to an NP which is inaccessible under the scope constraint on anaphoric

CHAPTER 2. THE EMPIRICAL VICINITY

relations and the sentential scope constraint.” (Roberts (1996a), p. 220)

That the ‘general case’ cannot be considered to be fully general becomes clear when Roberts notes that “the operator in the first sentence needn’t be intensional, as we see in (17) [my (23), D.B.], where it is an auxiliary negation” (Roberts (1996a), p. 220):

(23) John doesn’t have a car. It would be in the garage.

That this datum follows the scheme of Veltman’s puzzle is obvious, and at least the intuition that the example comes with an implicit antecedent is shared by Roberts. She writes: “The felicitous reading of the second sentence of (17) [my (23), D.B.] might be paraphrased ‘if John had a car, it would be in the garage.’” (Roberts (1996a), p. 221) However, she fails to elaborate on the issue in more detail. That this is a deficit is clear, when recalling the discussion of this kind of data that was given in chapter 1.¹⁰ Although negation does play a role in the interpretation of the data by supplying a negation induced alternative, the discussion has shown that negation can *not* (contrary to Roberts’ view quoted above) be held responsible for triggering modal subordination, if one is to use Roberts’ terminology here. This understanding is contradicted by examples like

(24) John repairs a car. It would not pass the TÜV approval test.

that lack an auxiliary negation in the initial discourse, but that are nonetheless felicitous and contain an implicit, contextually determined antecedent. The link to the context does in these cases arguably not come with an explicit intensional operator, but it is facilitated by the logical structure of the premise set required by

¹⁰An additional problem is that Roberts ignores the possible influence of the subjunctive mood in the example.

the subsequent discourse (i.e. the counterfactual consequent).^{11,12}

2.2 Contextually determined restrictions: principles and constraints

Having discussed the main questions regarding anaphora, an important open question that Roberts further addresses in her exposition of the modal and presuppositional data is the question of how to actually account for the peculiar interaction of context and intensional elements that is characteristic of the data. What is the relation between the parts of discourse involved? Or, more precisely, how can one account for the observed contextually restricted interpretation of intensional elements? Or, as Roberts puts it: “[T]he question which arises is what factors govern this relativization of the domain of the second intensional operator to a proposition suggested by the first sentence?” (Roberts (1996a), p. 224) –Which is of course a question also to be posed in connection with the discourses that yield Veltman’s puzzle.

Answering this question Roberts elaborates “[t]wo types of accounts appear to

¹¹ While (24) underlies a disjunctive premise set (the question under discussion connected to the counterfactual consequent being *When would the car not pass the TÜV approval test?*, while the answer is *When already a single precluding condition obtains*, the conjunctive (i.e. the positive) variant to the example is – independent from the occurrence of a negation in the initial sentence and as predicted by the proposal sketched in the preceding chapter – not-felicitous:

(25) #John (does not) repair(s) a car. It would pass the TÜV approval test.

The QUD here is *When do you pass the TÜV approval test?* and the answer is of course *When all conditions that are required are met at once*, i.e. when they conjunctively obtain.

¹² Actually, when shifting the perspective a bit, my proposal regarding the analysis of Veltman’s puzzle and Roberts’ general account may be reconciled to some degree. Assuming that the subjunctive mood acts as a form of (a grammaticalized) intensional operator the logical structure of the premise set does not so much ‘facilitate’ the link to the context – as I just claimed –, but it in fact just constrains the possibility of modal subordination. That the subjunctive introduces an intensional semantic element is beyond doubt. Indeed, considering the particular construction under discussion one might even argue that Roberts’ account of the ‘general case’ in terms of presuppositions that was quoted above might be applicable. Since in the examples that are of concern here the implicit antecedent of the counterfactual is presupposed to be false, while its complement has been asserted to be true in the discourse initial sentence, both build an *a priori* link.

CHAPTER 2. THE EMPIRICAL VICINITY

have the greatest explanatory potential, an entailment based, and hence semantic account, and one which crucially incorporates pragmatic processes, specifically accommodation.” (Roberts (1996a), p. 224) Anticipating the result of her discussion she eventually points out that “a purely semantic account is inadequate.” ((Roberts 1996a), p. 224) Yet, she continues: “[...] the two types of accounts aren’t really incompatible; assuming that entailment plays a role in licensing the anaphora in some examples does not preclude the necessity of using accommodation in many others. And in fact both have something to offer to our understanding of modal subordination; the lexical entailments of predicates naturally interact with other, contextual factors that constrain the accommodation that is possible in a given example.” ((Roberts 1996a), p. 224) In the next section I will first present the main tenets of Roberts’ discussion of the semantic account and will then follow up with a summary of her exploration of the pragmatic approach.

2.2.1 The semantic account in terms of entailments

At the heart of the semantic account of the data’s acceptability lies the concept of entailment. The idea is basically that felicity is due to the embedding of an expression into an entailing context. Roberts illustrates this via a special set of data. These examples are peculiar insofar as they – unlike most of the data discussed so far – do not involve any implicitly given semantic elements. Indeed, they exemplify the general scheme of a contextually given domain restriction that is at stake in an elementary and rather obvious way, but without any simplification. The restrictions derive immediately from the semantic relations of the involved intensional predicates. The pair of examples in (26) and (27) (first discussed in Lakoff (1970) and McCawley (1981)) demonstrates the point clearly and is further significant in that it exemplifies how entailments may determine the rift between felicitous and infelicitous discourses (cf. Roberts (1996a), p. 224; bold font and underlining as in the original):

(26) You are **required** to find a bear and **permitted** to take its picture.

(27) #You are **permitted** to find a bear and **required** to take its picture.

CHAPTER 2. THE EMPIRICAL VICINITY

Depending on their succession the lexical meanings – more specifically, the different deontic forces – of the verbs *permit* and *require* determine whether a discourse is felicitous or not.

Put loosely, the semantic intuition connected to (26) and (27) is that a permission does not commit the subject of the examples to choose a world in which she takes the picture of a bear. The picture may or may not be taken. Actually, a permission does not even presuppose the existence of a bear in any of the worlds accessible. Contrary with obligations. Being required to take a bear's picture does not leave the alternative not to, and indeed seems to presuppose the existence of bears that could be photographed. Applied to the data, the requirement expressed in the first example confronts one solely with bear worlds, in which – due to the second sentence – one may further take or not take a picture of a bear. In the second example however the domain specified by the initial sentence is (unlike in the first example) one in which also worlds are allowed in which bears don't even exist. Following up in discourse with the demand to take a picture of a bear in such a world (due to the universal force of the normative predicate) is then contradictory. For this reason the second example is infelicitous, while the first example is not: the *permission*-worlds are a subset of the *require*-worlds, but not the other way around.¹³ Thus, it is clear how entailments may determine discourse felicity.

Another example that nicely demonstrates the impact of lexical meanings and their relations on discourse felicity is given by the next data. While the infelicity of the first example (example (28)) is characterized by the same semantic preclusion as in example (27) above (with the difference that the scenario here is doxastic, rather than deontic), the simple change of the first predicate in (29) – and thus of the contextual background of the subsequent intensional predicate! – generates a felicitous discourse. Unlike in the first example the verbs involved in (29) describe a compatible semantic relation (cf. Roberts (1996a), p. 225 for (28), and p. 223

¹³ Note that – as also Roberts points out (cf. Roberts (1996a), footnote 8) – there is a second, felicitous reading of example (27) available in which the normative force of the obligation loses its universal appeal but is localized and restricted solely to those worlds in which one happens to find a bear. In this rendering the obligation comes with the qualification that only if one happens to find a bear, one is required to photograph it. (This point will find elaboration later on.)

CHAPTER 2. THE EMPIRICAL VICINITY

and p. 226 for (29); bold font and underlining again as in the original).

- (28) # Alice **denies** that there's a squirrel in her kitchen cabinets. She **hopes** to trap it alive and turn it loose outside.
- (29) Alice **fears** that there's a squirrel in her kitchen cabinets. She **hopes** to trap it alive and turn it loose outside.

The intuitive rationale that explains the difference in felicity in these examples is again based on the different modal forces connected to the verbs involved. All of them have a clear doxastic component: while *to deny* expresses a believe that negates the possibility that *p*, the verb *to hope* positively asserts the possibility that *p*. Thus, in consequence, the hope that *p* that is expressed in the second sentence of example (28) is interpreted before a background in which it already had been excluded that *p*. The semantic model generated is thus outright contradictory, which in turn explains again the infelicity of the discourse. Considering example (29) no such clash in modal forces arises: the fear entertained by the agent is a believe in a certain (unedifying) possibility. The existential import given by this finds concord in the existential supposition connected to the believe that translates from the agent's hope. (Note that however things are a little bit more complex than this rough rendering may suggest: the agent's *hope*-worlds coincide solely with those world's in which her fears are true, not with those in which her fears turn out to be false. Insofar does the initial sentence deliver a (localized) restriction of the domain of interpretation of the second sentence. In this the example runs parallel to the reading of example (27) mentioned in footnote 13. (The difference in the default readings of the examples in question is highly interesting, but cannot be pursued further at this point.)

Before turning to Roberts' critical assessment of the entailment approach, the perspective on the issue needs to be broadened further: subject to entailments are not only lexical meanings, but also presuppositions. With reference to Heim (1992) and Karttunen (1974) the matter is elaborated by Roberts at hand of data that are – as above – headed by propositional attitude verbs:

CHAPTER 2. THE EMPIRICAL VICINITY

(30) John believes that Mary is coming, and he wants Susan to come too.¹⁴

The point here is that the presupposition that is triggered by the particle *too* and that connects to the complement of *want* in the second conjunct, namely that *someone who is not Sue is coming*, is entailed by the first conjunct *John believes that Mary is coming*, because whenever John believes that Mary is coming, then in the worlds of concern it is true that someone who is not Sue is coming.

The assumption that “lack of entailment generally leads to infelicity” (Roberts 1996a:p. 227) – which follows from the postulate of a direct relation between entailments and discourse felicity as proposed in e.g. Lakoff (1970), Karttunen (1974), McCawley (1981), and Heim (1992), as Roberts points out – finds further support in example (31) (Roberts’ (21’)) in which the order of the intensional predicates involved in example (30) is inverted.

(31) #John wants Mary to come, although he believes that Susan is going to come too.

Again, the presupposition of the second conjunct is that someone else besides Susan is coming, however this presupposition is this time not satisfied by the initial sentence: wanting Mary to come does not guarantee that she will indeed do so. This lack of entailment hence explains the infelicity of the discourse.¹⁵

What these examples illustrate is that entailment relations and compatibilities between discourse elements resp. sentential constituents with different domains of interpretation influence felicity. (Indeed, the discourses are infelicitous not because some anaphoric constraint is violated, but because certain semantic relations between parts of discourse are incompatible.) Although Roberts does not bother to discuss elliptic cases, it seems plausible to assume that the same principles that govern the data just presented apply also here. After all, the mere fact that an el-

¹⁴My example (30) is Roberts’ (21), (cf. Roberts (1996a), p. 226).

¹⁵That Roberts does not address the change of the connective from *and* in (30) to *although* in (31) is problematic. My intuition is that (31) with *and* is felicitous. Which suggests that entailments (and thus felicity) are not determined solely by linearization, but that things are also mediated by the chosen connective.

CHAPTER 2. THE EMPIRICAL VICINITY

ement is given explicitly or implicitly arguably does not alter its presuppositions or domain of discourse. However, it is questionable, whether this account could be applied fruitfully to the analysis of Veltman's puzzle, because the entailment relations are in this case trivial.

2.2.2 Inadequacies of the semantic account – the pragmatic approach and the notion of *accommodation*

While there are strong reasons to back the semantic account, Roberts argues that although it is a sufficient condition, entailment represents nonetheless not a necessary condition for modal subordination. (cf. Roberts 1996a:p. 229). She points out that there are indeed a number of inadequacies to the entailment approach, all of which find a competitive solution in the concept of *accommodation*. In sum it is three arguments – none of which are independent from each other – that speak in Roberts' eyes against the semantic approach. The first argument highlights mismatches between predictions and actual judgments. Where the strictly logical perspective allows only for binary and categorial judgments and predicts infelicity due to non-entailment, actual judgments are often less prohibitive and appear more tolerant regarding felicity. Apparently, interpreters are able to charitably adjust domains of interpretation. The second argument is based on certain kinds of data that simply can not be accounted for in terms of entailments. And in the third argument Roberts brings up the concept of accommodation and its arguably superior explanatory value as an alternative to the entailment approach. Roberts puts her arguments as this: "First, many of the examples which an entailment-based account would predict to be bad seem to be less than fully infelicitous. Judgments tend to vary from informant to informant, variation which is typical of accommodation-based felicity (indeed predictable, given its nature), but not of entailment. Second, there are a number of types of examples which cannot be accounted for by entailment alone, and must involve the accommodation of the presuppositions in question; [...] And third, an alternative explanation of the infelicitous examples can be given within the context of an account involving accommodation." (Roberts (1996a), p. 227 - 228)

CHAPTER 2. THE EMPIRICAL VICINITY

Before addressing how Roberts substantiates these arguments, it is helpful to recall, as Roberts does (cf. Roberts (1996a), p. 215), the definition of accommodation, as given by David Lewis. Lewis writes:

If at time t something is said that requires presupposition P to be acceptable, and if P is not presupposed just before t , then – *ceteris paribus* and within certain limits – presupposition P comes into existence at t . (Lewis (1979b), p. 340.)

While the formulation of the first argument seems to invoke a criticism of the strict binary nature of entailment judgments, the point that Roberts wants to make is actually a different one. She argues that in some cases a secondary reading is available that supersedes judgments based on entailments and rather relies on accommodation as the base of interpretation. One such second reading connects to example (31) that had just been given above. The interpretation under which the sentence is felicitous can be made explicit as follows:

- (32) John wants Mary to come, although he believes that if she does, Susan is going to come too.

By accommodating to a best case scenario with respect to the *want*-worlds (insofar as the domain of the subject's beliefs is *ad hoc* restricted solely to those worlds in which the subject's wants have become reality) the entailment is rescued, and hence a felicitous reading is achieved. This is so, because if you just consider the worlds in which Mary actually shows up, then the presupposition that someone who is not Sue comes is met. What is important to be noted here is that accommodation is strictly local: it does not apply indiscriminately to all of the subject's *believe*-worlds – which would be global accommodation –, but it pertains solely to the subject's *make believe*-worlds. In other words, the *believe*-worlds are restricted and subordinated to the *want*-worlds.

Although the possibility of accommodation in examples like the above seems to affirm the importance of entailment in an indirect way, the picture shifts when turning to the details of the second argument. While the semantic approach unquestionably sheds an important light on the data, entailments can not be regarded

CHAPTER 2. THE EMPIRICAL VICINITY

to be the only determinative, as Roberts demonstrates.

The first example (due to Heim (1992)) that Roberts presents in her discussion of the second argument against an account of discourse felicity exclusively in terms of entailments appears to be actually quite similar to the ones just discussed. Also in (33) below (cf. Roberts (1996a), p. 229, her example (25)) two intensional predicates combine:

(33) John wants Fred to come, and he wants Jim to come too.

Checking for entailments is to no avail. John's *want*-worlds are such that either Fred or Jim or (ideally) both show up. But, as before, the semantics of *want* does not warrant that any of them actually shows up. Thus, the presupposition of the second conjunct (that is again triggered by the particle *too*) that *someone who is not Jim comes* is not entailed by the first conjunct. However, and this is an important difference to the above pair of examples (31) and (32), the example is felicitous from the start for any competent interpreter (Roberts does at least not make a remark to the contrary). Its felicity is arguably not based – as was the case in (31), resp. (32) – on a secondary reading that some interpreters may entertain, while others do not. Rather, the domain of the second occurrence of *want* is automatically restricted by the first *want* and its propositional complement. Which seems actually quite plausible, since both share the same domain. However, the entailment based approach would have a problem accounting for the felicity of the example. The problem actually extends, as Roberts continues to point out (cf. Roberts (1996a), p. 229), also to a plethora of other constructions, such as (to name just a few of what Roberts lists) adverbs of quantification, cases of mixed combinations of modal expressions, constructions containing attitude verbs in the first sentence which are followed by subjunctive propositions, and Geach's Hob/Nob-sentences (cf. Geach (1967)). And also, as Roberts mentions, examples “where the operator in the first sentence is negation”. (Roberts (1996a), p. 229). –Which brings us back to the actual concern. Clearly, entailments do not generally play a role here. Just consider cases where a nomological relation between the discourse elements is not presupposed, as for instance in the discourse

CHAPTER 2. THE EMPIRICAL VICINITY

John did not buy a newspaper. He would have missed the train. However, based on the discussion of the data, I do not follow Roberts' claim that negation is the decisive factor in the examples. In my eyes, it is indeed questionable whether the phenomenon under discussion really is a case of modal subordination in Roberts' sense. (But see again my footnote 12 on p. 48 for a brief discussion of this topic.) However, far more important than the 'proper' terminological classification of the phenomenon is to deliver a satisfactory analysis of the data at issue. A first, promising approach has briefly been sketched at the end of the preceding chapter. In the next chapter I will introduce the formal framework in which the analysis will eventually be spelled out. Yet, before doing so Schueler (2008) shall be briefly reviewed for its close empirical proximity to the data that I am concerned with.

2.3 Schueler (2008)

The data and concern of Schueler happens to be almost identical to mine: implicit counterfactual conditionals (which he abbreviates *ICs*). The question that also he pursues is "the problem of how to 'fill in' the antecedent when we are interpreting an *IC*". (Schueler (2008), p. 2) Schueler illustrates his point of departure by the following examples (cf. Schueler (2008), p. 12, examples (13a-c)):

- (34) John would kick a unicorn.
- (35) A yeti would ravage our town.
- (36) John would hate a war.

While these examples are tensed differently than mine (Schueler's data are in the modal present, while mine are in the modal perfect) the decisive difference to the data that I am discussing is something else. It is a rather serious disagreement in the conception and understanding of the data. Schueler insists that the data above represent felicitous sentences and are interpretable as they stand, i.e. in isolation and outside of any supporting context. Contrary to this I study solely (the discourse behavior and semantics of) counterfactual consequents within a context.

CHAPTER 2. THE EMPIRICAL VICINITY

The source for the interpretability of examples (34), (35), and (36) lies, so Schueler, with the (asserted and presuppositional) meaning of the indefinite NP. This sets examples of this kind apart from the data in (37), (38), and (39) below (= Schueler's examples (6a-c), p. 7), that do not contain an indefinite NP. In these cases the "understood antecedent" (Schueler (2008), p. 2) is not determined by the grammatical and semantic structure of what is overtly expressed in the consequent clause, but has to be derived from the context.

- (37) John would hate Paris.
- (38) Mary would kick John.
- (39) Sigmund would water the plant.

The distinction between what he calls "sentence-internal" (exemplified by (34), (35), and (36)) and "sentence-external" (exemplified by (37), (38), and (39)) implicit conditionals (cf. Schueler (2008), p. 3) yields the following two propositions which are to account for the implicit antecedent:

Sentence-External Antecedent-Filling: The antecedent in the correct paraphrase for the IC will be an antecedent which expresses the most plausible alterations to current contextual knowledge which would make the content of the IC (interpreted as a consequent in the paraphrase) relevant to the context of utterance. (Schueler (2008), p. 8)

Sentence-Internal Antecedent-Filling: The most contextually neutral interpretation for an IC containing an indefinite noun phrase is an interpretation best paraphrased by a full conditional where the indefinite noun phrase appears in the antecedent, and a corresponding pronoun appears in the consequent. (Schueler (2008), p. 15)

Focusing in his dissertation on "those [implicit conditionals] for which part of the understood antecedent is derived from the structure of the sentence" (Schueler (2008), p. 3), i.e. 'sentence-internal' implicit conditionals, the proposal he develops rests mainly on syntactic arguments (resp. arguments involving the interface of syntax and semantics, logical form). As the main tenet of his account

CHAPTER 2. THE EMPIRICAL VICINITY

of where the implicit antecedent comes from Schueler proposes ‘sideward covert movement’ (i.e. “movement to a non-c-commanding position” (Schueler (2008), p. 16)), taking its departure from the position of the explicit indefinite NP in the given consequent. I will not discuss Schueler’s syntactic analysis, since syntax lies not at the heart of interest of the present work. But some brief and general remarks about Schueler’s semantic conception shall be formulated, since I take this aspect of the account to be problematic in some respects.

Although I do not agree with Schueler’s intuition that the data he discusses are felicitous and interpretable as they stand, the conception can be motivated as follows. Given that the antecedent that is to complement the consequent expresses in any case a hypothetical proposition, Schueler’s idea is arguably the following: whenever the consequent contains hypothetical information (in form of an indefinite NP that introduces an actually non-existent, hypothetical referent into the discourse) this information – which is genuinely new information insofar as it comes without any presuppositional element, as is the case with e.g. definite NPs or proper names – may act as a natural filler for the (assumed and hypothetical) antecedent. –There is so to say a categorial match between part of what is given and what is to be assumed. Hence, the former stands in for the latter. This is the gist of the following quote. Schueler argues with respect to the examples (34) - (36):

These sentences differ from sentence-external ICs, in that ICs with indefinites can be interpreted with almost no context whatsoever. A natural interpretation of such sentences puts the indefinite at the heart of the understood antecedent. In case of (13[a-c]) [my (34) - (36), D.B.], since we have chosen noun phrases which introduce entities which don’t exist in our world (unicorns and yetis), it is specifically the hypothetical *existence* of something with the properties of the indefinite which forms the antecedent, with perhaps other properties attached. (In the case of war, although wars exist, their duration is limited and so there are many contexts in which the existence of a relevant war would be counterfactual as well.) That is, natural para-

CHAPTER 2. THE EMPIRICAL VICINITY

phrases of (13) [my (34) - (36), D.B.] are as in (14).

- (14)a. If a unicorn existed (and were in John's vicinity), John would kick it.
- b. If a yeti existed (and were in the vicinity of our town), it would ravage our town.
- c. If a war happened, John would hate it.

(Schueler (2008), p. 12; *italics* in the original)

While there may be some kind of categorial match between part of what is given and what is to be assumed, I do not follow Schueler that the former thus stringently stands in for the latter. I consider his understanding of the origin of the antecedent in case of the examples as too restricted. Contrary to what he suggests, the antecedent of a conditional is *not* determined by the grammatical structure or semantics of the consequent. Rather, one is *free* in what to assume and to express. The issue can be brought to the point at hand of his elaboration of the *war*-example to which he ascribes a certain counterfactual quality. But strictly speaking, this counterfactual element is inherent to virtually any possible referent. Not just *wars* are limited in their duration, but also for instance *parties* are so. But this does not compel one to interpret

(40) John would hate a party.

as

(41) If a party occurred, John would hate it.

rather than

(42) If there were no beer, John would hate a party.

Although in the latter case the lack of an appropriate antecedent would render (unlike as is the case with (41)) the introduction of an anaphoric pronoun impossible, this is not essential to making the assumption expressed in the antecedent. Thus,

CHAPTER 2. THE EMPIRICAL VICINITY

why should the consequent *John would hate a war* determine *If a war occurred, John would hate a war* rather than *If John were a pacifist, he would hate a war*? Certainly, the name *John* comes with the presupposition that there is a bearer of the name, but one might just as well assume that this bearer is a hypothetical, counterfactual individual, just as unicorns or yetis are.¹⁶

Indeed, when leaving the realm of events and entering the realm of objects, a similar case can be made up for instance for *bouquets* or *sandwiches* vs. *unicorns* and *yetis*. Since the existence of the former two is a rather limited existence one might think of “many contexts in which the existence of a relevant [such entity, D.B.] would be counterfactual as well” to re-quote Schueler (op. cit.). Following Schueler the generalization would thus be that

(43) John would buy a bouquet/a sandwich.

is ‘naturally’ understood as

(44) If a bouquet/a sandwich existed, John would buy it.

This however seems rather unintuitive. Of course, one might argue, that if one were to (be forced to) interpret these examples without any further context this might indeed be an opportune choice of an antecedent. But this is arguably a matter of *priming*, rather than a case of an assumption being determined by grammatical structure. Indeed, the lack of context does in my eyes not yield the “contextually most neutral” interpretation, as Schueler claims in the proposition concerned with *Sentence-Internal Antecedent Filling* quoted above. Rather than it being the most neutral context it is the most biased and determinate context, since if the “natural” understanding falls back on what the available context delivers and not on something else, this seems to just underline the importance of context for the reconstruction of the implicit antecedent. Apparently, interpreters take whatever the context delivers in order to accommodate the implicit antecedent.¹⁷

¹⁶That the indefinite NPs *a unicorn* and *a yeti* also have a generic reading is a problem that I don’t want to get into, but that Schueler does not go into in any mentionable detail either.

¹⁷I am using the term *accommodation* loosely here, not necessarily in the sense of Lewis.

CHAPTER 2. THE EMPIRICAL VICINITY

If this were not the case, and grammatical structure more important than context, then discourses like

(45) John is not a pacifist. He would hate a war.

were to be understood to contain the antecedent

(46) John is not a pacifist. If a war occurred, he would hate it.

Rather than:

(47) John is not a pacifist. If he were one, he would hate a war.

Finally, it seems that Schueler would have a problem to account for cases like the following. Given that the consequent *John would see a unicorn* has as its implicit antecedent *If a unicorn existed*, Schueler would arguably have a problem to account for why (48) is infelicitous, although its explicit counterpart (49) is:

(48) # John does not believe (a) unicorn(s) exist(s). He would see a unicorn.

(49) John does not believe (a) unicorn(s) exist(s). If a unicorn existed, he would see it.

2.4 Summary

The preceding chapter set out to locate the data under discussion within its empirical vicinity and to thus complete the picture. Most prominent in this vicinity is the range of data representative of the phenomenon of *modal subordination*. The role that contextually given information plays for the interpretation of subsequent parts of discourse provides with the link between the topic of modal subordination and Veltman's puzzle. In both cases intensional operators are given an implicit restriction by the preceding linguistic context, the biggest difference arguably being that in case of the data constituting Veltman's puzzle the operator connects to the expressed subjunctive mood, while in the classical modal subordination data the operator in question is a lexical expression. Thus, it seems plausible to subsume also the discourses under concern here under the concept of modal subordination,

CHAPTER 2. THE EMPIRICAL VICINITY

as for instance Roberts (1996a) does, based on a discourse like *John does not have a car. It would be in the garage (op. cit.)*, for which she claims the trigger is provided by the auxiliary negation in the first sentence. The impression however becomes problematic once the range of examples is extended, and also discourses as presented in the first chapter are considered. That Roberts' concept of modal subordination does not really apply – at least when understood as anaphora in intensional contexts, as Roberts also proposes – is best demonstrated by examples like # *John did not have surgery. He would not have died*. This case is infelicitous, even though the initial sentence contains a negation and thus provides with a restriction of the universally quantifying subjunctive conditional (in form of an implicit antecedent given by what I call a NinA) that would allow to resolve the reference of the pronoun in the counterfactual consequent.¹⁸ Further, none of the discussed semantic and pragmatic constraints on modal subordination (keywords are: entailments vs. accommodation) can explain the infelicity of the data. (If one is to maintain the idea that the data under discussion represent cases of modal subordination, then a third kind of constraint would have to be formulated.) Thus, the issues is not as clear as initially suggested.

Similarly for the account developed in Schueler (2008). Although also he discusses implicit counterfactual conditionals and asks where the implicit antecedent comes from it turns out upon closer investigation that data and understanding differ rather drastically from the examples I am discussing. While in the case of the discourses I am concerned with the antecedent is provided by the a context that precedes the consequent clause, Schueler restricts the attention of his study to cases of consequent clauses for which he argues that they are interpretable in isolation, and the antecedent being determined by information (originating from an indefinite NP) given in the consequent. The restriction is here thus provided by a semantic element occurring within the scope of the single conditional sentence to which the explicitly given information is supposed to act as the consequent. This is arguably also the reason why Schueler mentions modal subordination only in

¹⁸While the exact logical form of the counterfactual conditional has as of yet not been discussed, I take it to be obvious that it will involve in any case some universally quantifying intensional operator.

CHAPTER 2. THE EMPIRICAL VICINITY

passing in his discussion of the data. This ‘sentence internal’ analysis is orthogonal to the data I am concerned with and the ‘sentence external’ account of it that is to be proposed.

Chapter 3

The Formal Framework

Outline of the chapter

While the first chapter introduced Veltman's puzzle and presented the intricacies of the data, and the second chapter gave a broad overview of related phenomena and some of their theoretical aspects, in the next chapter the formal framework in which the analysis will be formulated is spelled out. The point of reference for this is provided again by [Veltman \(2005\)](#). After a brief introduction in section 3.1 to Veltman's actual concern (which differs from the one followed here), the chapter divides into three main parts. The first one, section 3.2, gives an informal motivation of the formal concepts to be introduced. The discussion elaborates in depth on certain arguments formulated in [Tichý \(1976\)](#). The upshot of this section is the insight that the interpretation of counterfactual conditionals is not simply a kind of minimal belief revision. Rather, it depends essentially on the recognition of the facts and laws that provides the frame of reference for the counterfactual. Ultimately expressing an inference (in the widest sense of the word), the logic and the premises on which the counterfactual rests are to be considered to determine the conditional's soundness. After these preparatory paragraphs the formal definitions of Veltman's semantics are introduced and discussed in section 3.3. Each definition is supported by an exposition of its import and meaning. Finally, to get practically acquainted with the workings of the technical apparatus and to understand its theoretical prowess, the formalism is implemented in the third main part of the chapter (section 3.4). It delivers a formal recount of the informal arguments

discussed earlier in part 3.2.

3.1 Locating Veltman's concern

3.1.1 How complex is the process of making counterfactual assumptions?

Apparently, the sequence of sentences that constitutes the discourse *John did not drink any wine. (If he had) He would have become sick* provokes to conduct something like a thought experiment: having incorporated the information that *John did not drink* given in the initial sentence, the subsequent counterfactual appeals to forget about this fact and to assume instead that circumstances were the opposite, i.e. to assume that *John drank*. The claim connected to this is that given this hypothetical state of affairs, then – *ceteris paribus* – it follows that *John would have become sick*. Thus, to make a counterfactual assumption $\lceil \text{If } \varphi \text{ had been the case, } \psi \text{ would have been the case} \rceil$ is a form of a three-step belief revision resp. deductive process that involves, firstly, the (momentarily) suspension of some factual information *not* φ in favor of, secondly, some hypothetical – indeed counterfactual – update of the information state with the proposition that φ , and, third, the claim that a certain consequence ψ logically derives from this change in premises from factually *not* φ to counterfactually φ .

While in principle the contours of this dialectic are rather simple, the question that Veltman pursues in his paper, namely how complex these steps actually are, turns out to be anything but simple. In fact the issue of counterfactual conditionals has generated a vast amount of literature, among the most influential being the accounts formulated in [Stalnaker \(1968\)](#) and [Lewis \(1973b\)](#). Most prominent among the problems posed by counterfactuals is the fact that counterfactual conditionals seem to defy a strict definition of their truth conditions. (For a classical discussion of the issue see for instance [Goodman \(1947\)](#).)

3.1.2 The answer in a nutshell

Anticipating already the main tenet of Veltman's analysis the answer is that it does not suffice to conceive of the process of making a counterfactual assumption \lceil *If it had been the case that φ* \rceil simply as a form of minimal belief revision that involves just the suspension of the factual information that *not φ* in favor of the counterfactual assumption that φ . Rather, the focus has to be widened, away from bare *facts* towards *relations between facts*, or, put differently, to premises *and* conclusions, and not just premises alone. Counterfactuals are rendered acceptable only in case they cohere with the laws and premises already assumed in the background of the counterfactual. The context and its logic has to be observed. The more so, since it delivers formally and conceptually the very backbone for the assessment of the validity of the counterfactual.

3.2 Understanding counterfactuals: an informal motivation of their formal interpretation

3.2.1 The Ramsey-test and the shortcomings of a naive belief revision account

The shortcomings of a simple belief revision analysis are discussed at length by Veltman. Point of departure for his elaboration is the so-called *Ramsey-test* (cf. Ramsey (1929)) that describes a general strategy for the evaluation of conditional statements. In a formulation due to Stalnaker (that differs slightly from Ramsey's original proposal, as Veltman points out, (cf. Veltman (2005), p. 162)) the Ramsey-test suggests the following heuristic for assessing (counterfactual) conditionals:

First, add the antecedent (hypothetically) to your stock of beliefs; second, make whatever adjustments are required to maintain consistency (without modifying the hypothetical belief in the antecedent); finally,

CHAPTER 3. THE FORMAL FRAMEWORK

consider whether or not the consequent is then true.

(Stalnaker (1968), p. 102)

When making a counterfactual assumption $\lceil \text{If } \varphi \text{ had been the case} \rceil$ the bare minimum required to maintain consistency of information is of course to suspend for the moment the information that in fact *not* φ . Thus, one might read the passage “*make whatever adjustments are required to maintain consistency*” as a directive for minimal change only. However, reducing the process of making a counterfactual assumption to the substitution of factual information by its counterfactual complement (while all other information is preserved) has its pitfalls, as we will see. It is more than the momentary retraction of factual information in favor of a provisional update with its counterfactual counterpart. Against the idea that to make a counterfactual assumption is to make only a minimal belief revision Veltman brings up a counterexample due to Pavel Tichý (formulated in Tichý (1976)).

3.2.2 An argument against the Ramsey-test: Tichý 1

Tichý’s counterexample is detailed in the following story (quoted from Veltman (2005), p. 162 - 163). The gist is that in making counterfactual assumptions not only the mere facts that are to be revised are involved, but – and this is decisive – also the rules that govern the relations between facts need to be recognized.

Consider a man, call him Jones, who is possessed of the following dispositions as regards wearing his hat. Bad weather invariably induces him to wear his hat. Fine weather, on the other hand, affects him neither way: on fine days he puts his hat on or leaves it on the peg, completely at random. Suppose moreover that actually the weather is bad, so Jones is wearing his hat.

The question that this example brings up is, in Veltman’s words: “[W]ould you accept the sentence ‘*If the weather had been fine, Jones would have been wearing his hat*’?” (Veltman (2005), p. 163)

While intuitively the answer is *no*, the account in terms of a simple minimal belief revision following Ramsey’s heuristic leads to a *yes* answer. To understand why

CHAPTER 3. THE FORMAL FRAMEWORK

this is so it is helpful to concentrate on the scenario in a bit more detail. Semi-formally restated, the two rules that determine the possible states of affairs are:

- rule 1: *weather bad* \rightarrow *hat on* 1
 rule 2: *weather good* \rightarrow *hat randomly on/off*

Given these two rules three worlds are available. One in which the weather is bad and Jones wears his hat. One in which Jones again wears his hat, but this time the weather is fine. And finally another one in which the weather is fine, but this time Jones does not wear his hat. The following table collects this information, with columns giving a world's attributes, while lines determine actual values²:

	<i>weather</i>	<i>hat</i>	
w_0	bad	on	as defined by rule 1
w_1	fine	on	as defined by rule 2
w_2	fine	off	as defined by rule 2

Since it is assumed that the weather is indeed bad, world w_1 is the actual world $w_@$.

If one is to follow the minimal belief revision story for counterfactuals, the only information that would have to be changed is the *weather*-parameter from *bad* to *fine*, while consistency considerations do *not* require a change in the *hat*-parameter: the hat remaining *on* is in accord with the condition specified for fine weather, since after all in fine weather the agent may chose to wear or not to wear the hat at random.

¹While with rule 1 Jones's wearing his hat is conditional (i.e. *dependent*) on the weather being bad, rule 2 is a comparably twisted conditional: it determines his behaviour regarding head dresses to be *independent* of the weather being fine. What is conditioned by the weather being fine is the randomness of Jones's behaviour, but not whether the hat is actually worn or not.

²In this table it already becomes clear how information is organized. Each world is determined by *facts* and *relations between facts*. Relations between facts are expressed by the rules that govern a scenario. Since for any scenario the possible facts happen to be either contingencies, or are determined by rules, it is useful – as Veltman does – to speak of the collection of basic information of this kind as the universe U_S of a state. The universe represents the logical space available. Alternatively to this I will occasionally use the notion of a 'horizon' to describe what possible worlds are accessible within a given information state.

CHAPTER 3. THE FORMAL FRAMEWORK

But why then tells intuition a different story? Why are we not ready to accept the counterfactual *If the weather had been fine, Jones would have been wearing his hat* as correct? The answer is because the minimal belief revision account ignores an almost trivial, but crucial factor: the facts that are specified in the scenario are not independent from the rules that constitute the scenario. While the hat being on is consistent with the hat being worn at random, it is *not* consistent with the background rules in the scenario that determine Jones's behaviour: in fine weather his habits concerning head dresses are *indeterminate*, while the counterfactual describes it as *determinate*. Thus, in this respect the counterfactual disagrees with the given information about the scenario.³ Hence, the minimal belief revision account leads to a counterintuitive result. (Actually, this is a rather important point. Information states are quite inert creatures when it comes to counterfactual deliberation. At least with respect to the nomological backbone of a scenario. Rules are persistent and tend to remain valid, thus delivering the frame of reference within which any revision has to take place.)

3.2.3 Tichý 1 as an argument against Stalnaker/Lewis

The same line of argument applies also to more explicitly minimalistic accounts of the meaning of (counterfactual) conditionals, such as the Stalnaker/Lewis approach, as formulated in Stalnaker (1968), and Lewis (1973b). (Which is the actual aim of Tichý's criticism, as Veltman points out.⁴) In these accounts, so Veltman, the truth condition for counterfactuals is given as follows:

A sentence of the form \ulcorner *If it had been the case that φ , it would have been the case that ψ* \urcorner is true in the actual world w iff the consequent ψ is true in every world⁵ in which the antecedent φ is true, and which

³The universal element that becomes clear in this is a rather dominant feature in the semantics of the counterfactuals that are of concern here. It will remain an important issue throughout the rest of the discussion.

⁴Cf. Veltman (2005), p. 163, where he writes: "Tichý's criticism was not directed directly against the Ramsey test but against the analysis of counterfactuals developed by Robert Stalnaker and David Lewis. (Stalnaker (1968), Lewis (1973b))".

⁵Veltman notes at this point "[a]ccording to Stalnaker there is at most one such world", while "according to Lewis there may be more than one." (Veltman (2005), p. 163, footnote 9.)

CHAPTER 3. THE FORMAL FRAMEWORK

in other respects differs minimally from w .

(Veltman (2005), p. 163)

However, also the conception of minimal difference that found entrance in the above truth condition gives rise to non-intuitive results for Tichý's scenario. (Counterfactually) assuming that the weather is fine, the world resembling the actual world w_0 the most is the one in which Jones is – just as in w_0 – wearing his hat. Thus, w_1 is 'closer' to w_0 than w_2 , since it differs only in the 'weather' value from w_0 , while w_2 differs in the values given for 'weather' *and* 'hat'.

3.2.4 Interim conclusion

The rift between theoretical account and intuitive understanding that emerges in these considerations highlights the importance of at least two points that are crucial in the interpretation of counterfactual conditionals: first, consistency is indeed pivotal when making counterfactual assumptions, and, second, the domain of consistency spans not only bare facts, but also – and decisively so – the rules and laws that organize these facts.

Note that also Lewis recognizes the importance of laws when it comes to maintaining consistency and reconciling counterfactual assumption with fact. Veltman quotes Lewis, who remarks: "(1) It is of the first importance to avoid big, widespread, diverse violations of law. (2) It is of the second importance to maximize the spatio-temporal region throughout which perfect match of particular fact prevails. (3) It is of the third importance to avoid even small, localized, simple violations of law." (Lewis (1979a), p. 472) However, Lewis insists at the same time that, fourthly, "[i]t is of little or no importance to secure approximate similarity of particular fact, even in matters that concern us greatly." (Lewis (1979a), p. 472) Taking this position, Tichý's example will of course lose its force, since within the fine weather worlds – i.e. every world in which the antecedent of the counterfactual is true – it appears as really nothing more than what Lewis calls a 'particular fact' – a contingency – whether the hat is worn or not. (Recall that in these worlds Jones decides at random about his headgear!) Hence, following

CHAPTER 3. THE FORMAL FRAMEWORK

Lewis, one is entitled to simply ignore this particular decision. This view, however, is problematic for more than one reason. When pondering next upon what might actually designate the concept of a ‘particular fact’ I will deviate from Veltman, but we will see, that a more detailed discussion is worthwhile: the notion of a ‘particular fact’ will turn out to be indeed untenable.

First, while locally, i.e. in the fine weather worlds, the hat wearing facts may indeed be ‘particular’ in the sense that they occur arbitrarily (i.e. independent from other facts), the picture changes, when a global perspective is taken and the set of all possible worlds is considered. It may appear paradoxical, but globally the (in)dependency of the *hat*-facts from the weather is actually dependent on the weather. After all, the rules that govern the scenario have *universal* appeal: only in case the weather is fine is Jones’s behaviour random. But in case the weather is bad, it is not. So, what counts as a ‘particular’ is dependent on the frame of reference. In the fine weather-worlds wearing a hat is certainly random, but that does not render it globally a particular. This however is not reflected in the Stalnaker/Lewis truth condition, and one may furthermore argue that the concept of a ‘particular fact’ is thus ill defined. Indeed, this is a point in principle, as we will see next. The following variant of the original example (henceforth referred to as *Tichý 1*, and *Tichý 2* respectively) drives the point a bit further.⁶ While in *Tichý 1* only consequences appear to be accidental, in *Tichý 2* also the premises are randomized, eventually revealing the full extend of the problem under discussion.⁷

3.2.5 Tichý 2: more against Lewis’s position

In *Tichý 2* Jones’s habits are described by Veltman as follows:

Suppose that Jones always flips a coin before he opens the curtains

⁶Veltman credits Frank Mulken, a former student of his, for conceiving the example.

⁷The issue actually goes deep into the question whether it is feasible to distinguish *rules* from *facts*. Unfortunately this question can not be addressed here in any more detail. Note however that, while *Tichý 2* is more explicit on the issue, the matter is already inherent to *Tichý 1*: since the weather being good or bad is a move of nature, Lewis allows one to simply ignore it when choosing among the counterfactual worlds. Hence, one simply looks at the world that is closest to the actual world. And that happens to be w_2 , independent of any condition that motivates the hat being worn or not.

CHAPTER 3. THE FORMAL FRAMEWORK

to see what the weather is like. Head means he is going to wear his hat in case the weather is fine, whereas tails means he is not going to wear his hat in that case. Like above, bad weather invariably makes him wear his hat. Now suppose that today heads came up when he flipped the coin, and that it is raining. So, again, Jones is wearing his hat.

(Veltman (2005), p. 164)

This amounts to the following abbreviation⁸:

- rule 1: $head \rightarrow (weather\ bad \rightarrow hat\ on)$
- rule 2: $tails \rightarrow (weather\ bad \rightarrow hat\ on)$
- rule 3: $head \rightarrow (weather\ fine \rightarrow hat\ on)$
- rule 4: $tails \rightarrow (weather\ fine \rightarrow hat\ off)$

Some consideration helps to get a clear idea of the intricacies of the scenario. While the rules are such that in some situations Jones's behavior is determined by a single parameter alone, in others the rules enforce that Jones's behavior is determined by a certain (inter-)dependency of these factors. The simple cases are (i) whenever *heads* comes up, and (ii) whenever the weather is *bad*. In both these cases each condition alone suffices to trigger Jones's choice of wearing a hat or not. Heads coming up means that Jones will wear his hat, no matter what the weather conditions are like, while, on the other side, the weather being bad determines that the outcome of the preceding coin toss is negligible: Jones will in any case wear his hat whenever the weather turns out to be bad. In case of tails coming up or the weather being fine, things are a bit more complex. Neither tails nor the weather being fine alone suffice to infer whether Jones will wear a hat or not. In these cases only the combination of both parameters determine Jones's behavior.

The next table summarizes the above information and collects all worlds that are in compliance with the rules that govern the scenario.

⁸Bracketing does not play a role here. When it comes later on in section 3.4.2, p. 103 to providing a proper translation I will use the way more elegant formalization employed by Veltman (cf. Veltman (2005), p. 170)

CHAPTER 3. THE FORMAL FRAMEWORK

	<i>coin</i>	<i>weather</i>	<i>hat</i>	
w_0	head	bad	on	= $w_@$
w_1	tails	bad	on	
w_2	head	fine	on	
w_3	tails	fine	off	

Consider next the case given in the scenario above. The actual world is distinguished by the coin toss having resulted in heads and the weather being bad. Hence, in accordance with rule 1, Jones is wearing his hat. Thus, $w_@$ is represented by w_0 .

Given this scenario and the rules of Tichý 2 the counterfactual *If the weather had been fine, Jones would have been wearing his hat* is intuitively true. What prompts this intuition shall be contemplated upon again in the following. This will not only serve to reconstruct the underlying understanding in more detail than Veltman bothers to do, but will also shed light on what is problematic about Lewis's opinion regarding the irrelevance of 'particular' fact. The findings turn out to be rather instructive.

Three factors play a major role in the line of argument, the first of which is rather trivial: (i) the semantic content of the counterfactual itself, (ii) the principle of inertia that underlies the idea of similarity that is employed in the evaluation of counterfactuals, and (iii) the randomness of the initial condition, i.e. the coin toss. Points (i) and (ii) go hand in hand and will be instrumental when explicating the underlying intuition, while point (iii) will – rather unsurprisingly – become crucial in the criticism of Lewis's position.

Thus prepared consider first the effects of the information conveyed by the counterfactual *If the weather had been fine, Jones would have been wearing his hat*. Given that w_0 is the actual world, the one parameter that the (antecedent of) the counterfactual intuitively commits one to change is the weather parameter from *bad* to *fine*. This immediately excludes w_1 from the set of counterfactual alternatives to the actual world w_0 , and leaves only worlds w_2 and w_3 up for consideration.

CHAPTER 3. THE FORMAL FRAMEWORK

However, since the counterfactual puts forth a *ceteris paribus* condition on the set of alternatives to $w_{@}$ in that it asks to consider only those worlds that differ from the actual world with respect to the conditional's antecedent, while everything else remains equal, only one of the two *fine weather*-worlds represents a viable counterpart to w_0 for the assessment of the counterfactual's truth. This world is w_2 , since w_3 differs (unlike w_2) not just in the *weather*-parameter from w_0 , but also in the outcome of the coin toss. For w_2 it then has to be checked whether the claim connected to the counterfactual consequent is true. And indeed, w_2 is a world in which Jones happens to wear a hat. The counterfactual conditional *If the weather had been fine, Jones would have been wearing his hat* is thus true in the given context. It describes a world that is consistent with the rules governing the scenario. (The most important for this particular case being rule 3 that states that whenever the coin landed heads and the weather turned out to be fine, Jones chooses to wear a hat.)

After retracing the intuitive understanding, the next question is whether one can achieve the same result following Lewis's recipe? The answer is no. Given Lewis's statement that "[i]t is of little or no importance to secure approximate similarity of particular fact" (*op cit.*) one may simply ignore the result of the coin toss and disregard it, since it represents nothing more than a contingent fact in the scenario. But if particulars of this kind were really negligible, one might just as well opt for the coin having landed tails. Under these circumstances however by rule 4 it follows that Jones would not have worn a hat. The counterfactual would thus have to be regarded to be false. Which is not the intuitive interpretation. What this shows is that, unlike Lewis suggests, you can not ignore contingent premises in assessing the truth or falsity of counterfactual conditionals.

Before I will turn in the next section to the formal foundations of the framework with which Veltman tackles the matter under discussion, first his assessment of the topic shall be quoted. While it may differ slightly in its perspective from my discussion of the issue, the theme is nonetheless the same. He summarizes the problem as follows:

CHAPTER 3. THE FORMAL FRAMEWORK

What really matters is this: In both cases [i.e. Tichý 1 and Tichý 2, D.B.] Jones is wearing his hat *because* the weather is bad. In both cases we have to give up the proposition that the weather is bad – the very *reason* why Jones is wearing his hat. So, why should we want to keep assuming that he has his hat on? In the first case there is no special reason to do so; hence, we do not. In the second case there is a special reason. We will keep assuming that Jones is wearing his hat because we do not want to give up the independent information that the coin came down heads. And this, together with the counterfactual assumption that the weather is fine, brings in its train that Jones would have been wearing his hat. In other words, similarity of particular fact is important, but only for facts that do not depend on other facts. Facts stand and fall together. [Footnote omitted, D.B.] In making a counterfactual assumption, we are prepared to give up everything that depends on something that we must give up to maintain consistency. But we want to keep in as many independent facts as we can. (Veltman (2005), p. 164 - 165; *italics* in the original)

This concludes the informal discussion of what it means to make counterfactual assumptions. The technical challenges that come along with this should have gained some contour by now. Next, the formalism developed by Veltman to tackle these issues shall be introduced in detail.

3.3 The formalism

3.3.1 Some definitions of elementary concepts: *update*, *support*, and *entailment*

The framework is an update semantics with a cognitive twist. As characteristic for all dynamic semantics, meaning is modelled in this framework in terms of context change potential: the meaning of a linguistic expression lies in its capacity to change the information state entertained by a linguistic agent. However, Veltman conceives of this basic idea in more broader, cognitive terms, by identifying

CHAPTER 3. THE FORMAL FRAMEWORK

information states with cognitive states. This is technically emulated as follows:

- The meaning $[\varphi]$ of a sentence φ is an operation on cognitive states.

In the following ‘ $S[\varphi]$ ’ denotes the result of applying the operation $[\varphi]$ to state S ; it is the result of *updating* S with φ .

(Veltman (2005), p. 160)

Next, Veltman determines how cognitive states cohere. The notion of *support* is most elementary in this respect:

A cognitive state S supports a sentence φ when updating S with φ adds no information over and above what is already in S . Instead of ‘ S supports φ ’, I will often say ‘ φ is accepted in S ’.

- S supports a sentence φ , $S \models \varphi$, iff $S[\varphi] = S$.

(Veltman (2005) p. 160)⁹

Since the system can not end up with ‘more’ information than it started out with the idea that interpretation is a reduction of possibilities (in the available logical space of a state) seems rather natural. The discussion of Tichý 1 and Tichý 2 already suggested this. It also suggested to think of the procedural side of the problem of interpretation in terms of a derivational process. However, since none of these issues are elaborated by Veltman in detail, I will not speculate any further in this direction. But note that these considerations would eventually take over to Veltman’s definition of *logical validity* that follows next. It is defined in terms of the concept of *support*:

⁹With this the conceptual and formal landscape of Veltman’s framework is decisively constrained. Apparently, it is not possible to incorporate *genuinely new* information. Rather, any update is already inherent to the agent’s cognitive state. This seems to commit one to a rather holistic idea of meaning and information: You can only understand what you can actually grasp, given the knowledge that you have. Premises thus arguably play a rather important role when it comes to interpretation. This will find some reverberation later on in Definition 5, where Veltman defines counterfactuals as tests, and comments: “Given this definition, sentences of the form ‘*if had been φ , would have been ψ* ’ do not convey new information – not directly at least.” (Veltman (2005), p. 171)

CHAPTER 3. THE FORMAL FRAMEWORK

[A] sequence of premises $\varphi_1, \dots, \varphi_n$ entails a conclusion ψ if updating a state with that sequence invariably leads to a state that supports the conclusion. (Veltman (2005), p. 161)

Formally this is rendered as follows:

- $\varphi_1, \dots, \varphi_n \models \psi$ iff for any state S , $S[\varphi_1] \dots [\varphi_n] \models \psi$.
(Veltman (2005), p. 161)¹⁰

In the following definitions Veltman specifies the technical machinery necessary to provide an account for counterfactual updates that avoids the pitfalls addressed in the introductory section.

3.3.2 The system's ontology

Presupposing, as already Veltman does, basic knowledge of the concepts of possible world semantics, I will address next with Definition 1 the first cornerstone of his framework. It defines the system's ontology.

Definition 1 (Worlds and states)

Fix a finite set \mathcal{A} of atomic sentences.

- (i) A *world* is a function with domain \mathcal{A} and range $\{0, 1\}$; a *situation* is a partial such function; a *proposition* is a set of worlds.
- (ii) Let W be the set of possible worlds. A *cognitive state* S is a pair $\langle U_S, F_S \rangle$, where either (a) $\emptyset \neq F_S \subseteq U_S \subseteq W$; or (b) $F_S = U_S = \emptyset$.

(Veltman (2005), p. 165)

¹⁰Following this conception it would be interesting to endow Veltman's system with a corresponding deduction theorem. A dynamic semantic framework in which model-theoretic and proof-theoretic ideas go hand in hand is formulated in van Lambalgen and Hamm (2005). The formulation of a hybrid between Veltman's theory and van Lambalgen and Hamm's system seems not too far off.

CHAPTER 3. THE FORMAL FRAMEWORK

Before elaborating on clause (ii) and Veltman’s two-dimensional notion of the concept of a cognitive state I will first briefly review clause (i). Note that while Veltman states that “[i]n this definition a possible world is identified with the valuation that assigns the value 1 to the atomic sentences true in it, and 0 to the atomic sentences false in it.” (Veltman (2005), p. 165), it might be added that – assuming that the set of sentences \mathcal{A} is just large enough – worlds can be considered to be full descriptions of the states of affairs. Any fragment of this description consequently defines a *situation*. The short hand Veltman uses to designate the proposition φ is the usual $\llbracket\varphi\rrbracket$. I will adopt this convention throughout the discussion.

For the basic inventory of propositional logic a standard set-theoretical interpretation is assumed. The following definitions apply (cf. Veltman (2005), p. 165):¹¹

$$\begin{aligned} \llbracket p \rrbracket &= \{w \in W \mid w(p) = 1\} \\ \llbracket \neg\varphi \rrbracket &= W \sim \llbracket \varphi \rrbracket \\ \llbracket \varphi \wedge \psi \rrbracket &= \llbracket \varphi \rrbracket \cap \llbracket \psi \rrbracket \\ \llbracket \varphi \vee \psi \rrbracket &= \llbracket \varphi \rrbracket \cup \llbracket \psi \rrbracket \\ \llbracket \varphi \rightarrow \psi \rrbracket &= (W \sim \llbracket \varphi \rrbracket) \cup \llbracket \psi \rrbracket \end{aligned}$$

More interesting than these necessary provisions is Veltman’s definition of a cognitive state as an ordered pair as given in clause (ii) of Definition 1. Veltman separates here what might be called *rule base* and *data base*: while U_S determines the *universe* of a state by collecting not only the facts, but also the rules that define relations between facts (the meta-facts so to say), in F_S only the facts are collected, but not the rules. Thus the intuitive distinction made in the Tichý-examples finds a reflex in the very foundation of Veltman’s framework. Indeed, this idea to separate rules and facts (or, more precisely, the (nomologically) possible and the factual) lies at the heart of his proposal. Note the use of the empty

¹¹While the first line makes up for the missing condition on set comprehension in the mentioning of the notion of a *proposition* in Definition 1, the conception of a proposition here unfortunately defies a straightforward translation into a type-driven semantics. This is one major drawback of Veltman’s theory. One advantage however is that it allows for the definition of the notion of a *situation* (as given in Definition 1(i)) that will be decisive later on. (For another approach to counterfactuals that also employs *situations* (in a more canonical situations-theoretic semantics) see Kratzer (1989).)

CHAPTER 3. THE FORMAL FRAMEWORK

set in clause (ii b) of Definition 1. It represents the *absurd* (state): there simply is no world that could model this state. Veltman comments on this: “Agents will avoid getting into this state.” (Veltman (2005), p. 166) To keep things simple he allows for only *one* absurd state (given as $\mathbf{0} = \langle \emptyset, \emptyset \rangle$) by identifying all such states (cf. Veltman (2005), p. 166). Orthogonal to the absurd state is the so called *minimal state* in which no world is excluded. The minimal state represents a state of total ignorance, being ‘minimal’ in the sense that no information whatsoever (be that nomological, or factual) about the world is given, and hence nothing can be precluded. Consequently, it is represented by $\mathbf{1} = \langle W, W \rangle$. Finally, a state is regarded to be *coherent*, whenever it is not contradictory, i.e. not absurd: $F_S \neq \emptyset$. (cf. Veltman (2005), p. 166)

3.3.3 The kernel of Veltman’s update semantics

The effect that the meaning of an expression φ has on a state, when the state is updated with φ is modeled in Definition 2 below. Updates are differentiated by the two basic types of information that are distinguished in this framework: mere factual statements versus law-like expressions. Each of these effectuate change in the respective domains of a cognitive state.¹²

Law-like propositions are prefixed by a box operator ‘ \square ’. The box operator indicates their nomological status and designates the expression to be read as ‘*it is a law that ...*’, with the dots “to be filled by a formula of propositional logic”, as Veltman remarks (cf. Veltman (2005), p. 166 - 167). Clause (ii) of Definition 2 takes care of this, while clause (i) deals with factual updates.

Definition 2 (Interpretation)

Let φ be a formula of propositional logic.

- (i) (a) $S[\varphi] = \langle U_S, F_S \cap \llbracket \varphi \rrbracket \rangle$ if $F_S \cap \llbracket \varphi \rrbracket \neq \emptyset$;
- (b) $S[\varphi] = \mathbf{0}$, otherwise.
- (ii) (a) $S[\square\varphi] = \langle U_S \cap \llbracket \varphi \rrbracket, F_S \cap \llbracket \varphi \rrbracket \rangle$ if $F_S \cap \llbracket \varphi \rrbracket \neq \emptyset$;

¹²To be more precise, since $F_S \subseteq U_S$, updates of U_S trivially affect also F_S .

CHAPTER 3. THE FORMAL FRAMEWORK

(b) $S[\Box\varphi] = \mathbf{0}$, otherwise.

(Veltman (2005), p. 167)

Incorporating new (factual) information in form of some non-boxed proposition φ is determined as the result of updating the state S with the information that φ . The result of an update with the proposition φ is eliminative in the sense that with the update that φ the $\neg\varphi$ worlds are discarded. This is modeled by intersecting F_S and φ (Provided, that, of course, the intersection of F_S and φ is not empty. In case the intersection is empty, the worlds described by φ are not F_S worlds, i.e. φ is not a statement about anything possible with regards to the given facts F_S . Hence one reaches the absurd state.)

Nomological updates inevitably relate to U_S since a new rule is to be incorporated into the universe. Again, the update is deemed absurd, if the rule does not comply with the factual information, i.e. if the intersection of F_S and φ is empty.

Veltman summarizes this as follows:

Updating with a propositional formula φ eliminates from F_S all possible worlds in which φ is false. Hence, only worlds in which φ is true are left as worlds that might be the actual world. If there are no such worlds left, one gets into the absurd state. Similarly, an update with $\Box\varphi$ eliminates from U_S all worlds in which φ is false. So, only worlds in which φ is true are left as worlds that might have been the actual world. The other ones are so outlandish, you do not have to reckon with them, not even in making the wildest counterfactual assumption.
(Veltman (2005), p. 167)

(In a footnote to this he remarks: “Note that this definition would not work if we had allowed stacking of \Box ’s etc.” (Veltman (2005), p. 167, footnote 14), the reason being that in the two-dimensional definition of cognitive states laws about laws can not be represented.)

To illustrate the workings of the definitions introduced so far consider a language $\mathcal{A}_{p,q,r}$ with just three atoms p , q , and r . The world(s) describable by this fragment

CHAPTER 3. THE FORMAL FRAMEWORK

could obtain eight different constellations of facts in form of $w_0 - w_7$. Assume next a (with respect to $\mathcal{A}_{p,q,r}$) linguistically fully competent agent, who does not know anything about the world she is in. In other words, the agent is in the minimal state **1**, and anything expressible by $\mathcal{A}_{p,q,r}$ could (still) be the case. According to Definition 1 ff. the agent's cognitive state S is thus:

⟨	U_S	p	q	r	,	F_S	p	q	r	⟩
	w_0	0	0	0		w_0	0	0	0	
	w_1	0	0	1		w_1	0	0	1	
	w_2	0	1	0		w_2	0	1	0	
	w_3	0	1	1		w_3	0	1	1	
	w_4	1	0	0		w_4	1	0	0	
	w_5	1	0	1		w_5	1	0	1	
	w_6	1	1	0		w_6	1	1	0	
	w_7	1	1	1		w_7	1	1	1	

Assume that the first bit of linguistic information about the world witnessed by the agent is given by the law that *whenever q, then r*, i.e. $\Box(q \rightarrow r)$. This nomological update is prohibitive in the sense that it stipulates that it can not be the case that q is true, while at the same time r is false. In effect the law excludes certain worlds from ever having to be considered as possible, and determines not just what is or is not the case, but what can not be the case.¹³ As such nomological updates affect not just the ontic dimension of the state, i.e. F_S , but also the ontological dimension in form of the universe U_S , or, in Veltman's words: "[T]he general laws set a limit to the factual information one can have." (Veltman (2005), p. 166) Definition 2, clause (ii) captures the domain-transcending, universal character of the box-operator and gives a precise formulation of the point. Instantiated to the present example the following picture emerges:

Since the minimal state **1** is defined as $\langle W, W \rangle$, and the meaning of the conditional *If q, then r* ($= \llbracket q \rightarrow r \rrbracket$) is rendered classically to be $(W \sim \llbracket q \rrbracket) \cup \llbracket r \rrbracket$, the

¹³This is of course a highly problematic simplification. Although an interesting point, I will however not attempt a more subtle differentiation of the semantics of laws and law-like expressions in relation to the different modal domains.

CHAPTER 3. THE FORMAL FRAMEWORK

nomological update of the minimal state **1** with the information that $\Box(q \rightarrow r)$ yields the following altered state (in which it still holds that $U_S = F_S$):

$$\mathbf{1}[\Box(q \rightarrow r)] = \langle U_S \cap \llbracket \Box(q \rightarrow r) \rrbracket, F_S \cap \llbracket \Box(q \rightarrow r) \rrbracket \rangle$$

Further computation shows that the update is not absurd, since $F_S \cap \llbracket \Box(q \rightarrow r) \rrbracket \neq \emptyset$. (Propositions are presented in due course by the set of worlds in which the respective sentence is true. Since in the state under discussion $U_S \cap \llbracket \Box(q \rightarrow r) \rrbracket = F_S \cap \llbracket \Box(q \rightarrow r) \rrbracket$ I will for reasons of readability just compute the first coordinate of the ordered pair in detail. In the last line the result will eventually be copied to the second coordinate, too.)

$$\begin{aligned} & U_S \cap \llbracket \Box(q \rightarrow r) \rrbracket \\ &= U_S \cap ((W \sim \llbracket q \rrbracket) \cup \llbracket r \rrbracket) \\ &= U_S \cap ((W \sim \{w_2, w_3, w_6, w_7\}) \cup \{w_1, w_3, w_5, w_7\}) \\ &= U_S \cap ((\{w_0, w_1, w_4, w_5\}) \cup \{w_1, w_3, w_5, w_7\}) \\ &= U_S \cap (\{w_0, w_1, w_3, w_4, w_5, w_7\}) \\ &= \{w_0, w_1, w_3, w_4, w_5, w_7\} \end{aligned}$$

Hence, the state $S = \mathbf{1}[\Box(q \rightarrow r)]$ is the tuple of the sets of worlds:

$$\langle \{w_0, w_1, w_3, w_4, w_5, w_7\}_{=U_S}, \{w_0, w_1, w_3, w_4, w_5, w_7\}_{=F_S} \rangle$$

The pair of tables below gives the state $\mathbf{1}[\Box(q \rightarrow r)]$ in a more accessible presentation. The worlds w_2 , and w_6 violate the given law and are thus discarded:

⟨	U_S		p		q		r
	w_0		0		0		0
	w_1		0		0		1
	w_2		/0/		/1/		/0/
	w_3		0		1		1
	w_4		1		0		0
	w_5		1		0		1
	w_6		/1/		/1/		/0/
	w_7		1		1		1

,

⟩	F_S		p		q		r
	w_0		0		0		0
	w_1		0		0		1
	w_2		/0/		/1/		/0/
	w_3		0		1		1
	w_4		1		0		0
	w_5		1		0		1
	w_6		/1/		/1/		/0/
	w_7		1		1		1

CHAPTER 3. THE FORMAL FRAMEWORK

Next consider a factual update of the state $\mathbf{1}[\Box(q \rightarrow r)]$ that was just achieved. Assume that the agent subsequently learns that in the world she is facing the fact that p and the fact that q are true. Incorporating this information into the cognitive state S is achieved by an update of S with the proposition expressed by the formula $p \wedge q$. The state to be reached is $S = \mathbf{1}[\Box(q \rightarrow r)][p \wedge q]$. Since the update is not a nomological update, just the factual dimension of S is affected. The meaning of $p \wedge q$ being given by the intersection of the sets of worlds in which p resp. q is true, the update yields according to Definition 2 clause (i) the following state (if $F_S \cap \llbracket p \wedge q \rrbracket \neq \emptyset$):

$$\begin{aligned}
 & \mathbf{1}[\Box(q \rightarrow r)][p \wedge q] \\
 &= \langle U_S, F_S \cap \llbracket p \wedge q \rrbracket \rangle \\
 &= \langle U_S, F_S \cap \llbracket p \rrbracket \cap \llbracket q \rrbracket \rangle \\
 &= \langle U_S, F_S \cap \{w \in W \mid w(p) = 1\} \cap \{w \in W \mid w(q) = 1\} \rangle
 \end{aligned}$$

That $F_S \cap \llbracket p \wedge q \rrbracket \neq \emptyset$ can be seen directly: since $\llbracket p \rrbracket = \{w_4, w_5, w_6, w_7\}$ and $\llbracket q \rrbracket = \{w_2, w_3, w_6, w_7\}$ and $F_S = \{w_0, w_1, w_3, w_4, w_5, w_7\}$ it follows that $F_S \cap \llbracket p \wedge q \rrbracket = \{w_7\}$.

This means that – as it turns out – the only world that is law-abiding and in which the conjunction of the facts p and q is true is world w_7 . All other worlds are discarded (or, better, de-actualized) in the factual dimension F_S of the state $\mathbf{1}[\Box(q \rightarrow r)][p \wedge q]$:

⟨	U_S	p	q	r	,	⟩
	w_0	0	0	0		
	w_1	0	0	1		
	w_2	/Ø	/N	/Ø		
	w_3	0	1	1		
	w_4	1	0	0		
	w_5	1	0	1		
	w_6	/N	/N	/Ø		
	w_7	1	1	1		
	F_S	p	q	r		
	w_0	/Ø	/Ø	/Ø		
	w_1	/Ø	/Ø	/N		
	w_2	/Ø	/N	/Ø		
	w_3	/Ø	/N	/N		
	w_4	/N	/Ø	/Ø		
	w_5	/N	/Ø	/N		
	w_6	/N	/N	/Ø		
	w_7	1	1	1		

CHAPTER 3. THE FORMAL FRAMEWORK

What happens if one tries to contradict the information that is constitutive of the state can be seen by trying to update the state $\mathbf{1}[\Box(q \rightarrow r)][p \wedge q]$ with the information that r is false, although the state designates it to be true. (This follows by *Modus Ponens* from the law $\Box(q \rightarrow r)$ and the fact that q having been established during the course of updates.) Since the worlds in which r is in fact not the case are $\{w_0, w_2, w_4, w_6\}$, while the state $\mathbf{1}[\Box(q \rightarrow r)][p \wedge q]$ is represented by $\{w_7\}$, the update $\mathbf{1}[\Box(q \rightarrow r)][p \wedge q][\neg r]$ results in the empty set, \emptyset . It is not representable in the logical space: being told that *not* r , while *knowing* that r just does not make any sense in terms of information transfer. Hence, as Veltman puts it, agents try to avoid reaching this state.¹⁴

3.3.4 What it means to make counterfactual assumptions: preliminaries

Before spelling out the formal account of what it means to make a counterfactual assumption, first some more groundwork needs to be laid. The preceding discussion of the Tichý-examples highlighted the importance of rules for counterfactual deliberation: any counterfactual update occurs within a context of premises in form of rules and laws. In Definition 3 below Veltman identifies certain properties pertaining to this domain that will become relevant later on. I will first present the definition and will then comment in loose prose upon its parts and the concepts introduced therein.

Definition 3 (Basis)

Let $S = \langle U_S, F_S \rangle$ be a state.

- (i) The situation s *forces* the proposition P *within* U_S iff for every $w \in U_S$ such that $s \subseteq w$ it holds that $w \in P$.¹⁵

¹⁴While the absurd state is not a goal in itself, ending up in it certainly does more good than harm, and is not necessarily to be avoided. At least if one perceives of updates as tests for the robustness and consistency of the established information. In case the test fails and one reaches the absurd state this just indicates that either the preceding state has to be revised, or that the assumed shared communicative background is deficient and has to be re-assessed in communication. Insofar does the absurd state play an essential role.

¹⁵In footnote 15, p. 168 Veltman adds at this point: “If there is no world $w \in U_S$ such that $s \subseteq w$, then, according to this definition, the situation s forces every proposition.”

CHAPTER 3. THE FORMAL FRAMEWORK

- (ii) The situation s *determines* the world w iff s forces $\{w\}$ within U_S .
- (iii) The situation s is a *basis* for the world w iff s is a *minimal* situation determining w within U_S .

(Veltman (2005), p. 167 - 168)

With the main part of clause (i) determining that some fragment s of the governing body of laws (and facts) can be considered to *force* a proposition P when given s also P is valid (i.e. s and P coincide) the notion of *forcing* characterizes a peculiar relation between premises. Put somewhat hyperbolically *forcing* reflects a possible dependency between premises in U_S , in the form of one premise coming in the wake of another premise (or set of premises), stating that every s -world is also a P -world. (Which thus arguably allows to recognize derivative facts, or, alternatively, premises for premises, and possible structure within U_S .) In clauses (ii) and (iii) the idea underlying the notion of *forcing* is escalated further. In (ii) the proposition forced by s is valid in only one world, namely the world of s . In other words, given the situation s there is no alternative left to the world of which s is a part of. Thus, the situation s indeed uniquely *determines* the world it is a part of. Finally, a *basis* as introduced in clause (iii) is the least amount of information necessary and sufficient to exactly identify a world in U_S . Or, in Veltman's words:

A basis for a world $w \in U_S$ is a part of w consisting of mutually independent facts which, given the general laws, bring the other facts constituting w in their train. (Veltman (2005), p. 168)

When discussing the formal analysis of the Tichý-examples later in more detail than Veltman does, the full scope of the introduced notions will become clearer. However, already now it shall be pointed out that the antecedent of a conditional rule that is element of U_S will force the consequent as soon as the antecedent is given as a fact. While I will later also discuss how bases are determined, note already now that when a conditional rule and the facticity of its antecedent are given, the antecedent acts as a *basis* of the rule's consequent. However, before further elaborating on these points I will first focus on the most important definition of Veltman's theory.

3.3.5 What it means to make counterfactual assumptions: the formal account

In the next definition, the core of Veltman's proposal is reached. Definition 4 below seeks to emulate what is involved when making a counterfactual assumption. In a nutshell the idea is to suspend the factual counterpart to whatever assumption the counterfactual introduces, while to simultaneously preserve and pay heed to whatever laws and facts remain constitutive of the situation. This eventually provides the ground for the counterfactual update without it falling subject to the shortcomings of the simple belief revision story. As exercised before I will first present the definition and will then follow up with an informal recount of the matter.

Definition 4 (Retraction)

Let $S = \langle U_S, F_S \rangle$ be a state.

- (i) Suppose $w \in U_S$, and $P \subseteq W$. The set $w \downarrow P$ is determined as follows: $s \in w \downarrow P$ iff $s \subseteq w$ and there is a basis s' for w such that s is a *maximal* subset of s' not forcing P .
- (ii) $S \downarrow P$, the *retraction of P from S* , is the state $\langle U_{S \downarrow P}, F_{S \downarrow P} \rangle$ determined as follows:
 - (a) $w \in U_{S \downarrow P}$ iff $w \in U_S$;
 - (b) $w \in F_{S \downarrow P}$ iff $w \in U_S$ and there are $w' \in F_S$ and $s \in w' \downarrow P$ such that $s \subseteq w$.
- (iii) The state S [*if it had been the case that φ*] is given by $(S \downarrow \llbracket \neg \varphi \rrbracket)[\varphi]$.

(Veltman (2005), p. 168)

Contemplating upon the import of this definition the first point to notice is that while clause (ii) directly elaborates on the effects the suspension of information brings about to a cognitive state, the core of the issue is already delivered in clause (i). Its concern lies with the objects constitutive of any state: *possible worlds*.

CHAPTER 3. THE FORMAL FRAMEWORK

Clause (i) elaborates on how to model the necessary change that a retraction of information evokes on each of these cognitive atoms. –After all, any global change of a cognitive state is due to local change in its underlying.

The first concern of clause (i) is to constrain the set of worlds available in the logical space. The worlds to consider are solely those that comply with the relevant information, i.e. only worlds within the universe U_S of the given state. With respect to (each of) these worlds some fact is to be suspended. This fact is given by the truth assignment for some proposition P at the particular world w , and the goal is to define w minus the fact that P . (The thus altered world – i.e. set of truth assignments – is given by $w \downarrow P$.) Now, the lesson to be learned from the Tichý-examples was that the world(s) to be achieved by retracting P must be such that they still comply with the premises of the situation. Recall that for instance in the case of Tichý 1 the problem was that following the simple belief revision account the resulting world(s) were determinate with respect to Jones’s behavior, while the scenario did not allow for this to be the case, but contained a rule that declared his behavior to be random. –Which in turn explained the unintuitiveness of the naive account. In order to prevent this and to get a more accurate result Veltman falls back on his concept of a *basis*. (Recall that a *basis* is the minimal amount of information that determines a world that complies with the laws and facts.) The idea formulated in the main part of clause (i) is to modify the bases of the worlds under consideration such that they do not force P anymore.¹⁶ This way it is guaranteed that the worlds defined are no longer P -worlds, but are still observant of the laws relevant in the scenario. The recipe that captures the essence of clause (i) is thus to take any world $w \in U_S$, to look at its basis s' , and to determine a maximal subset s of this basis such that this situation s does not force P anymore. This way a situation is determined that lies at the heart of those worlds that are no

¹⁶The subset s is to be maximal in order to guarantee maximal similarity of the counterfactual worlds to the actual world.

CHAPTER 3. THE FORMAL FRAMEWORK

longer P -worlds.^{17,18}

Building on this idea clause (ii) of Definition 4 elaborates the consequences of what has been formulated in clause (i) for the global concept of a cognitive state: Veltman defines here the change in the cognitive state that the suspension of information will yield. In the first half of clause (ii) – condition (a), that is concerned with the first coordinate of the ordered pair that represents a cognitive state – Veltman builds on the intuition that in any case the laws that had already been established will remain valid. Hence Veltman stipulates that any world constitutive of the universe of the altered state $U_{S\downarrow P}$ will be such that it is part of U_S , the universe of the state *ex ante*.¹⁹

Next follows condition (b) which elaborates on the second coordinate of the state: $F_{S\downarrow P}$ resp. its source F_S . What is detailed here is that in any case the altered facts (the worlds $w \in F_{S\downarrow P}$) have to comply with U_S and that they are different from

¹⁷ While I rather sloppily referred to the situation s as a basis for the *not P*-worlds, note that strictly speaking s is not a basis in the technical sense, since s might not determine just a single *not P*-world, but might be part of multiple *not P*-worlds. Note also that while Veltman appears to use the subset relation ' \subseteq ' it actually is a shorthand designating a partial function. Recall Definition 1, clause (i): "A *world* is a function with domain \mathcal{A} [\mathcal{A} being a finite set of atomic sentences, D.B.] and range $\{0, 1\}$; a *situation* is a **partial such function** [my emphasis, D.B.]; a *proposition* is a set of worlds."

¹⁸Note that by considering a *maximal* subset of the basis Veltman incorporates into the definition of retraction a similarity measure for counterfactual worlds with respect to the actual world(s). How intricate the issue is is brought to the point by David Lewis when discussing his famous example *If kangaross had no tails, they would topple over* (cf. Lewis (1973b), p. 1). Lewis writes: "We might think it best to confine our attention to worlds where kangaroos have no tails and *everything* else is as it actually is; but there are no such worlds. Are we to suppose that kangaroos have no tails but that their tracks in the sand are as they actually are? Then we shall have to suppose that these tracks are produced in a way quite different from the actual way. Are we to suppose that kangaroos have no tails but that their genetic makeup is as it actually is? Then we shall have to suppose that genes control growth in a way quite different from the actual way (or else that there is something, unlike anything there actually is, that removes the tails). And so it goes; respects of similarity and difference trade off. If we try too hard for exact similarity to the actual world in one respect, we will get excessive differences in some other respect." (Lewis (1973b), p. 9; *italics* in the original.)

¹⁹Thus conceived laws are made exempt from counterfactual deliberation. This represents a strong restriction of the analytic powers of Veltman's theory. While one would like to be able to treat laws, i.e. stipulated relations between facts, similar to the treatment of 'simple' facts when making counterfactual assumptions, the issue defies a clear solution in Veltman's theory. The point will prove to be problematic on several more occasions.

CHAPTER 3. THE FORMAL FRAMEWORK

the actually established facts F_S only insofar as one is no longer committed to assume that P . Or, sticking closer to the original formulation, he defines the worlds of the altered state $F_{S \downarrow P}$ to be such that they share some situation s with the factual worlds F_S that however does not commit one to the proposition P .²⁰

Finally, in clause (iii) a notational convention is fixed that determines how to refer to counterfactually updated states.

3.4 Excursion: putting theory to practice – computing Tichý’s examples in Veltman’s system

Having given a so far mostly abstract synopsis of the core of Veltman’s proposal I will next demonstrate its practical implementation by applying the formalism to the analysis of first Tichý 1, and then Tichý 2. The goal of this step-by-step exercise is to consolidate and to deepen the understanding of Veltman’s theory of counterfactuals.

However, first two minor remarks about how semantic intuitions and some of the discussed theoretical concepts relate shall be made. Since the points to be made will play a role in the following, I will address them beforehand.

The first of these remarks is concerned with the universal appeal of counterfactuals and the question whether a single witness suffices to establish a counterfactual truth. The intuition is that counterfactuals are universal in the sense that in all worlds described by the counterfactual it has to be true. (That is, the domain of the counterfactual needs to be exhausted in this respect.) In fact – and this relates intuition and Veltman’s theory – whatever the counterfactual worlds are, they have to have a basis such that the claim put forward by the counterfactual *determines* the worlds w under consideration within U_S .

Next, it shall be brought to mind again that this basis needs to be different from the

²⁰Again, this implies some similarity measure for the worlds to be considered.

CHAPTER 3. THE FORMAL FRAMEWORK

basis of those worlds that stand for the actual world $w_{@}$ only insofar as in these counterfactual worlds w the (factual) antecedent is assumed to be not true. Or, more precisely, it *must not be forced by the circumstances*. Everything else about these worlds ought to remain the same.

How-to determine a *basis*

Being already concerned with the concept of a *basis*, another major key to Frank Veltman’s proposal is to understand how to determine a *basis* for a world. Recall that the only hint that Veltman gives is provided by the following quote: “A basis for a world $w \in U_S$ is a part of w consisting of mutually independent facts which, given the general laws, bring the other facts constituting w in their train.” (Veltman (2005), p. 168) While I will not give a general algorithm for how to compute the basis of a given world, I will nonetheless sketch a general line of reasoning one might pursue. Having an idea of how to approach the matter should provide with helpful insight for solving the matter when actually facing the problem.²¹

To keep things simple assume that the language under consideration has only two atoms p and q , and that laws are always of the conditional form $\Box(\phi \rightarrow \psi)$.

Given such a two-atomic language the minimal state **1** is²²:

1	p	q
w_0	1	1
w_1	1	0
w_2	0	1
w_3	0	0

When pursuing the question of how to determine a basis, the task at hand seems at first to be rather complex. This is, however, not the case. Pivotal for a so-

²¹Talking of ‘*the*’ basis is a slight misnomer. We will see later, when discussing Tichý 2 (cf. page 107), that, as Veltman observes, a world might have not just a single basis: “[...] generally speaking, it may very well be that a world has more than one basis”. (Veltman (2005), p. 168)

²²In order to make things no more complex than necessary I will give a state as a one dimensional matrix instead of an ordered pair of $\langle U_S, F_S \rangle$. The point to be made will not suffer from this simplification.

CHAPTER 3. THE FORMAL FRAMEWORK

lution of the problem are the two notions of *general laws* and *independent facts* that Veltman alludes to in his informal rendering of the concept of a *basis* already cited above. With regards to the concept of an *independent fact* it is (at least for expository reasons²³) helpful to distinguish two types of facts: for one, those facts that figure as premises (axioms) in the derivation of other facts, and, second, those factual elements that do not stand in any relation to any other facts, but may be considered to be fully independent. While both kinds of facts are obviously important in the identification of a world, the former play a special role in the process of reasoning to other facts. This process will, in any case, involve whatever general laws have been established besides the facts. –Which brings the two notions of *facts* and *laws* that lie at the very center of the concept of a *basis* into the proper relation.

As long as no laws have been fixed, all given facts are independent from each other, and no fact suffices to determine or establish any other fact. Hence, for the minimal state **1** every world acts as its own base, i.e. the bases are given by the worlds themselves. The situation changes however once laws that determine relations between facts have been introduced. Assume for illustration that the minimal state **1** has been updated with a conditional law of the form $\Box(p \rightarrow q)$. Since this excludes all worlds in which the antecedent is true but the consequent is false, the set of worlds W is reduced to the universe U_S of this nomologically updated state. Hence, U_S (= the state $\mathbf{1}[\Box(p \rightarrow q)]$) consequently contains only those worlds in compliance with the law introduced in the state.²⁴ These worlds are w_0 , w_2 , and w_3 , while world w_1 is eliminated from further consideration in this state:

U_S	p	q
w_0	1	1
w_1	/1/	/0/
w_2	0	1
w_3	0	0

²³The distinction to be made is ultimately not tenable.

²⁴To mention it again, this is a problematic point. It ultimately might give rise to a position that assumes that what must not be the case, can not be the case.

CHAPTER 3. THE FORMAL FRAMEWORK

Next, assume that the fact p is indeed true.²⁵ From this it follows immediately that also q has to be true, since under the premises $\Box(p \rightarrow q)$ and p the conclusion that q can not be false (at least if one assumes (i) the validity of *Modus Ponens* as a universal rule of inference and (ii) one interprets the conditional as material implication). Conversely, from $\Box(p \rightarrow q)$ being true and q being false it follows by *Modus Tollens* that p must also be false, since q being false, but p being true would contradict what had been established as valid, namely the law that $\Box(p \rightarrow q)$ with which the minimal state had been updated.

Taking stock at this point, it follows from these considerations that in state $\mathbf{1}[\Box(p \rightarrow q)]$ the basis for w_0 is provided by the fact that p (abbreviated $\langle p, 1 \rangle$), while the basis of w_3 is given by $\langle q, 0 \rangle$: in combination with the law constitutive of U_S both facts entail so much information in terms of other facts that they eventually allow to determine their respective worlds as a whole. While in the cases of w_0 and w_3 the matter is rather straightforward, things are a bit more complicated than this with respect to world w_2 . In case of w_2 no single fact alone is sufficient enough to deduce the other fact. This is so because the premises $\Box(p \rightarrow q)$ and $\langle p, 1 \rangle$ are insofar *indeterminate* with respect to the world they entail, as they permit the possibility of $\langle q, 0 \rangle$ or $\langle q, 1 \rangle$. –Which leaves not just one possible world, but two possible worlds (namely w_2 and w_3) compatible with the given information. Alternatively assuming $\langle q, 1 \rangle$ in combination with the law $\Box(p \rightarrow q)$ yields a similarly inconclusive result. In this case again two possibilities arise that are in compliance with the law: $\langle p, 1 \rangle$ and $\langle p, 0 \rangle$, which ultimately leaves the worlds w_0 and w_2 as possible extensions. In light of this it turns out that the only information set that allows to uniquely determine w_2 is the conjunction of the given factual information. Hence, the basis for w_2 is provided by the set $\{\langle p, 0 \rangle, \langle q, 1 \rangle\}$.

When summarizing the above, two rules of thumb give its gist: facing a conditional law and a true possible instantiation of the law's antecedent, one can deduce by *Modus Ponens* the value of the consequent, since it must not be false. Hence,

²⁵While any such factual update will of course primarily alter F_S , recall that in Definition 3 bases are determined with respect to U_S . Hence, for the moment it suffices again to focus solely on U_S , and to ignore F_S completely.

CHAPTER 3. THE FORMAL FRAMEWORK

the antecedent acts as the basis. On the other hand, when facing a conditional law and a false possible instantiation of the law's consequent, by *Modus Tollens* one can derive that the antecedent must not be true, but false. Hence, in this case the consequent figures as the basis.

While Veltman does not bother to elaborate on this, note at least in an aside that of course the underlying logic and the rules of inference applied are decisive for the determination of a basis. Note further that each basis will also depend upon what general laws are in effect. Assuming for instance $\Box(p \rightarrow \neg q)$ instead of $\Box(p \rightarrow q)$ the basis for each world will be different from the ones just determined (though of course the line of reasoning effective will be basically the same): while in this case the law eliminates w_0 from U_S (since $\langle p, 1 \rangle$ and $\langle q, 1 \rangle$, which means $\langle \neg q, 0 \rangle$), the basis for w_1 is $\langle p, 1 \rangle$, the basis for w_2 is $\langle q, 1 \rangle$, and the basis for w_3 is $\{\langle p, 0 \rangle, \langle q, 1 \rangle\}$. Note next that a basis is also relative to the domain under consideration: Shifting from the minimal state to for instance a nomologically updated state will change the basis for any world, since the updated domain is 'smaller' than that of the minimal state itself. Finally, another observation that is intimately related to the last remark is that bases are *dynamic*. They *evolve* during the succession of updates. The basis that determines w_0 at the minimal state **1** is different from the basis of w_0 at state **1** [$\Box(p \rightarrow q)$], the former being given by the assignments $\langle p, 1 \rangle, \langle q, 1 \rangle$, the latter being $\langle p, 1 \rangle$.

After these preliminaries I will turn now to the long overdue step-by-step analysis of Tichý 1. First, the story line of Tichý 1 will receive a formal translation. Next, departing from the minimal state, the succession of the nomological and factual updates is formally executed. And finally, the attention will rest upon the practical details and intricacies connected to the assumed counterfactual update.

3.4.1 Tichý 1

The initial set-up: the rules, the facts, and the minimal state

Recollecting the material body of Tichý 1, the scenario is comprised of the following givens: on the one hand the general rules that characterize Jones's behavior,

CHAPTER 3. THE FORMAL FRAMEWORK

and on the other hand the facts that specify the particular situation we will look at. While the facts alone determine F_S , i.e. what is actually the case, the laws (in combination with the facts) define what could potentially be the case, i.e. U_S , the universe of the situation. Rendered to a quasi-formal presentation the rules are:

- rule 1: *weather bad* \rightarrow *hat on* 26
rule 2: *weather good* \rightarrow *hat randomly on/off*

The facts divide into those that are explicitly mentioned in the scenario, and those that are deducible from the given information. Directly given is:

- fact 1: *weather bad*

Indirectly given (following from *Fact 1* and *Rule 1*) is:

- fact 2: *hat on*

Assuming a strong complementary relation between the meaning of the sentences *The weather is bad* and *The weather is fine* the translation into a propositional language (that is endowed with the usual connectives) is straightforward. In the following the atom p translates the proposition *The weather is bad*, while the proposition *The weather is fine* is translated by $\neg p$. Finally, the proposition *Jones is wearing his hat* converts to q .

The minimal state **1** will thus consist of at least two atoms p and q . However, in order to achieve greater generality and to avoid being unduly restricted solely to the particular I will follow Veltman and enrich the space of possibilities by a third proposition r (cf. Veltman (2005), p. 169). While this extension is (at least at the moment²⁷) not really decisive, one avoids having to stipulate a certain technical patch in the case one faces empty intersections when determining the bases of counterfactual worlds. For more details on this see my footnote 35 on p. 100,

²⁶The use of the adverbial *randomly* is rather sloppy here. While the semantics of *randomly* allows for a translation using a bi-conditional, rule 2 is actually not compatible with this. The final formalization will pay heed to this point.

²⁷Later on, when turning to the analysis of the data that provides the actual starting point of this study, the introduction of the third propositional variable will turn out to be of great explanatory value.

CHAPTER 3. THE FORMAL FRAMEWORK

though.

Thus, the three propositions p , q , r delineate the logical space that defines the minimal state given by the ordered pair $\langle U_S, F_S \rangle$:

\langle	<table style="border-collapse: collapse; text-align: center;"> <thead> <tr> <th style="padding: 2px 5px;">U_S</th> <th style="padding: 2px 5px;">p</th> <th style="padding: 2px 5px;">q</th> <th style="padding: 2px 5px;">r</th> </tr> </thead> <tbody> <tr><td style="padding: 2px 5px;">w_0</td><td style="padding: 2px 5px;">0</td><td style="padding: 2px 5px;">0</td><td style="padding: 2px 5px;">0</td></tr> <tr><td style="padding: 2px 5px;">w_1</td><td style="padding: 2px 5px;">0</td><td style="padding: 2px 5px;">0</td><td style="padding: 2px 5px;">1</td></tr> <tr><td style="padding: 2px 5px;">w_2</td><td style="padding: 2px 5px;">0</td><td style="padding: 2px 5px;">1</td><td style="padding: 2px 5px;">0</td></tr> <tr><td style="padding: 2px 5px;">w_3</td><td style="padding: 2px 5px;">0</td><td style="padding: 2px 5px;">1</td><td style="padding: 2px 5px;">1</td></tr> <tr><td style="padding: 2px 5px;">w_4</td><td style="padding: 2px 5px;">1</td><td style="padding: 2px 5px;">0</td><td style="padding: 2px 5px;">0</td></tr> <tr><td style="padding: 2px 5px;">w_5</td><td style="padding: 2px 5px;">1</td><td style="padding: 2px 5px;">0</td><td style="padding: 2px 5px;">1</td></tr> <tr><td style="padding: 2px 5px;">w_6</td><td style="padding: 2px 5px;">1</td><td style="padding: 2px 5px;">1</td><td style="padding: 2px 5px;">0</td></tr> <tr><td style="padding: 2px 5px;">w_7</td><td style="padding: 2px 5px;">1</td><td style="padding: 2px 5px;">1</td><td style="padding: 2px 5px;">1</td></tr> </tbody> </table>	U_S	p	q	r	w_0	0	0	0	w_1	0	0	1	w_2	0	1	0	w_3	0	1	1	w_4	1	0	0	w_5	1	0	1	w_6	1	1	0	w_7	1	1	1	$,$	<table style="border-collapse: collapse; text-align: center;"> <thead> <tr> <th style="padding: 2px 5px;">F_S</th> <th style="padding: 2px 5px;">p</th> <th style="padding: 2px 5px;">q</th> <th style="padding: 2px 5px;">r</th> </tr> </thead> <tbody> <tr><td style="padding: 2px 5px;">w_0</td><td style="padding: 2px 5px;">0</td><td style="padding: 2px 5px;">0</td><td style="padding: 2px 5px;">0</td></tr> <tr><td style="padding: 2px 5px;">w_1</td><td style="padding: 2px 5px;">0</td><td style="padding: 2px 5px;">0</td><td style="padding: 2px 5px;">1</td></tr> <tr><td style="padding: 2px 5px;">w_2</td><td style="padding: 2px 5px;">0</td><td style="padding: 2px 5px;">1</td><td style="padding: 2px 5px;">0</td></tr> <tr><td style="padding: 2px 5px;">w_3</td><td style="padding: 2px 5px;">0</td><td style="padding: 2px 5px;">1</td><td style="padding: 2px 5px;">1</td></tr> <tr><td style="padding: 2px 5px;">w_4</td><td style="padding: 2px 5px;">1</td><td style="padding: 2px 5px;">0</td><td style="padding: 2px 5px;">0</td></tr> <tr><td style="padding: 2px 5px;">w_5</td><td style="padding: 2px 5px;">1</td><td style="padding: 2px 5px;">0</td><td style="padding: 2px 5px;">1</td></tr> <tr><td style="padding: 2px 5px;">w_6</td><td style="padding: 2px 5px;">1</td><td style="padding: 2px 5px;">1</td><td style="padding: 2px 5px;">0</td></tr> <tr><td style="padding: 2px 5px;">w_7</td><td style="padding: 2px 5px;">1</td><td style="padding: 2px 5px;">1</td><td style="padding: 2px 5px;">1</td></tr> </tbody> </table>	F_S	p	q	r	w_0	0	0	0	w_1	0	0	1	w_2	0	1	0	w_3	0	1	1	w_4	1	0	0	w_5	1	0	1	w_6	1	1	0	w_7	1	1	1	\rangle
U_S	p	q	r																																																																									
w_0	0	0	0																																																																									
w_1	0	0	1																																																																									
w_2	0	1	0																																																																									
w_3	0	1	1																																																																									
w_4	1	0	0																																																																									
w_5	1	0	1																																																																									
w_6	1	1	0																																																																									
w_7	1	1	1																																																																									
F_S	p	q	r																																																																									
w_0	0	0	0																																																																									
w_1	0	0	1																																																																									
w_2	0	1	0																																																																									
w_3	0	1	1																																																																									
w_4	1	0	0																																																																									
w_5	1	0	1																																																																									
w_6	1	1	0																																																																									
w_7	1	1	1																																																																									

This shall provide the starting point for the exercise to follow. In due course the matrix will be updated with the information that determines the scenario of Tichý 1.

The rules and the corresponding nomological updates

The first update is due to rule 1, i.e. the information that *If the weather is bad, Jones will wear his hat* being communicated. The translation of this general law is: $\Box(p \rightarrow q)$. This eliminates worlds w_4 and w_5 from the domain of consideration, since these worlds defy the validity of the law.²⁸ The state yielded by updating the minimal state $\mathbf{1}$ with $\Box(p \rightarrow q)$ is $\mathbf{1}[\Box(p \rightarrow q)]$:

²⁸It is worth to note again that nomological updates pertain to both coordinates of a state.

CHAPTER 3. THE FORMAL FRAMEWORK

⟨	U_S	p	q	r	,	F_S	p	q	r	⟩
	<i>w</i> ₀	0	0	0		<i>w</i> ₀	0	0	0	
	<i>w</i> ₁	0	0	1		<i>w</i> ₁	0	0	1	
	<i>w</i> ₂	0	1	0		<i>w</i> ₂	0	1	0	
	<i>w</i> ₃	0	1	1		<i>w</i> ₃	0	1	1	
	<i>w</i>₄	/N	/Q	/Q		<i>w</i>₄	/N	/Q	/Q	
	<i>w</i>₅	/N	/Q	/N		<i>w</i>₅	/N	/Q	/N	
	<i>w</i> ₆	1	1	0		<i>w</i> ₆	1	1	0	
	<i>w</i> ₇	1	1	1		<i>w</i> ₇	1	1	1	

Next follows an update with rule 2 which states that *If the weather is fine Jones will or will not wear a hat*. Since its propositional translation $\Box(\neg p \rightarrow q \vee \neg q)$ is a tautology, the preceding state $\mathbf{1}[\Box(p \rightarrow q)]$ and the resulting state $\mathbf{1}[\Box(p \rightarrow q)][\Box(\neg p \rightarrow q \vee \neg q)]$ are equivalent:

⟨	U_S	p	q	r	,	F_S	p	q	r	⟩
	<i>w</i> ₀	0	0	0		<i>w</i> ₀	0	0	0	
	<i>w</i> ₁	0	0	1		<i>w</i> ₁	0	0	1	
	<i>w</i> ₂	0	1	0		<i>w</i> ₂	0	1	0	
	<i>w</i> ₃	0	1	1		<i>w</i> ₃	0	1	1	
	<i>w</i>₄	/N	/Q	/Q		<i>w</i>₄	/N	/Q	/Q	
	<i>w</i>₅	/N	/Q	/N		<i>w</i>₅	/N	/Q	/N	
	<i>w</i> ₆	1	1	0		<i>w</i> ₆	1	1	0	
	<i>w</i> ₇	1	1	1		<i>w</i> ₇	1	1	1	

(Note that, since by now all the rules that govern the scenario of Tichý 1 have been successfully updated, the universe U_S of Tichý 1 has been fixed in this state.)

The facts and the factual updates

Next come the updates with factual information. Strictly following Tichý's story the only update concerns the information that the weather is indeed bad. Hence, the state $\mathbf{1}[\Box(p \rightarrow q)][\Box(\neg p \rightarrow q \vee \neg q)]$ is updated by the fact that p . This distinguishes the actual from the non-factual worlds, and is commonly considered

CHAPTER 3. THE FORMAL FRAMEWORK

to lead to the elimination of the *not p*-worlds.²⁹ (This assumption will however be suppressed in the table. Rather the *not p*-worlds are immediately designated as counterfactual worlds (indicated by the triple lines in the table).³⁰ Note that the division between the factual and the counterfactual applies only to the dimension of the actual, i.e. F_S , not to the dimension of the potential, i.e. U_S . The resulting state is $\mathbf{1}[\Box(p \rightarrow q)][\Box(\neg p \rightarrow q \vee \neg q)][p]$:

⟨	U_S	p	q	r	,	F_S	p	q	r	⟩
	w_0	0	0	0		w_0	0	0	0	
	w_1	0	0	1		w_1	0	0	1	
	w_2	0	1	0		w_2	0	1	0	
	w_3	0	1	1		w_3	0	1	1	
	w_4	/N	/Q	/Q		w_4	/N	/Q	/Q	
	w_5	/N	/Q	/N		w_5	/N	/Q	/N	
	w_6	1	1	0		w_6	1	1	0	
	w_7	1	1	1		w_7	1	1	1	

Note that I will not conceive of derivative information such as the fact that q (since p and rule 1 are given) to be generated in a separate update. Rather the notion of *support* covers such acts of implicit information transfer/generation. The question whether the order of updates is important (factual information before nomological information, or, as presented here, nomological information before factual information) shall not be pursued. In either case the state reached would be the same as the one depicted. What shall be emphasized however at this point of the presentation is that one can identify a basis for each of the worlds given in the state under discussion.³¹ The bases are listed below. For each of these bases it is spelled out

²⁹cf. for instance Veltman (2005), p. 167, where Veltman remarks: “Updating with a propositional formula φ eliminates from F_S all possible worlds in which φ is false. Hence, only worlds in which φ is true are left as worlds that might be the actual world.”

³⁰Actually, as I will take it, the notion of ‘elimination’ here is a misnomer. The *not p* worlds are not discarded. They do not simply disappear from the logical space resp. the agent’s cognitive state. Rather than to eliminate worlds in the strict sense negation evokes a *division* of the set of worlds. It distinguishes the factual from the counterfactual dimension. Thus, it would be more appropriate to speak of a *de-actualization*, rather than an *elimination*.

³¹Actually, bases can be determined at any point of the update process. The choice to do so now is due to expository reasons: the bases will become relevant in a moment when reaching the counterfactual update.

CHAPTER 3. THE FORMAL FRAMEWORK

how it determines its respective world by providing a derivation that exemplifies how the combination of factual and nomological premises allows to derive the ‘missing’ facts. This way each basis is indirectly proven to be sound. In addition to this also the utility of the concept of a *basis* and its interaction with the underlying logic is demonstrated. This provides more insight into the workings of Veltman’s framework than a mechanical recount of how the bases are identified could deliver.

The basis for w_6 is $\{\langle p, 1 \rangle, \langle r, 0 \rangle\}$. The ‘missing’ information about the world derives as follows:

$$\text{Modus Ponens} \frac{\langle \Box(p \rightarrow q), 1 \rangle \quad \langle p, 1 \rangle \quad \langle r, 0 \rangle}{\langle q, 1 \rangle}$$

The basis for w_7 is $\{\langle p, 1 \rangle, \langle r, 1 \rangle\}$. Its world is fully determined by:

$$\text{Modus Ponens} \frac{\langle \Box(p \rightarrow q), 1 \rangle \quad \langle p, 1 \rangle \quad \langle r, 1 \rangle}{\langle q, 1 \rangle}$$

The basis for the counterfactual worlds w_0 and w_1 is justifiable by *Modus Tollens*, using the specific combination of values for q and r :

The basis for w_0 is $\{\langle q, 0 \rangle, \langle r, 0 \rangle\}$. It derives:

$$\text{Modus Tollens} \frac{\langle \Box(p \rightarrow q), 1 \rangle \quad \langle q, 0 \rangle \quad \langle r, 0 \rangle}{\langle p, 0 \rangle}$$

The basis for w_1 is $\{\langle q, 0 \rangle, \langle r, 1 \rangle\}$. It allows to infer the value of the rule’s antecedent as the last bit of information to be known about the world.

$$\text{Modus Tollens} \frac{\langle \Box(p \rightarrow q), 1 \rangle \quad \langle q, 0 \rangle \quad \langle r, 1 \rangle}{\langle p, 0 \rangle}$$

For worlds w_2 and w_3 the combinatorics of truth values for p and q is such that they are mutually indeterminable given a classical logic. Hence, the bases of these worlds are given only by the full information available in these worlds: The basis for w_2 is $\{\langle p, 0 \rangle, \langle q, 1 \rangle, \langle r, 0 \rangle\}$. The basis for w_3 is $\{\langle p, 0 \rangle, \langle q, 1 \rangle, \langle r, 1 \rangle\}$.

CHAPTER 3. THE FORMAL FRAMEWORK

The counterfactual update(s)

While so far the nomological and the factual updates have been discussed, what remains to be examined is how the system accounts for counterfactual updates. This, of course, is the most interesting issue now, the question being whether the counterfactual *If the weather had been fine, Jones would have been wearing his hat* turns out to be true or not in the account. Recall that intuitively it is false. In the following I will go over how in Veltman's framework the issue is solved.

Some preparatory remarks: a rough reconstruction of how things work

Recapitulating the main insights of the discussion so far will help to get a clear idea of what will have to be done.

Whatever the counterfactual *fine weather worlds*-worlds are they have to have a basis such that

- the basis is informed by the laws in effect for the scenario
- the basis needs to be different from the basis of $w_{@}$ only insofar in these worlds the (factual) antecedent is assumed to be not true, or, more precisely, *must not be forced* by the circumstances

But how is p not forced? The answer to this immediate question will provide with the next step in the consideration. So, recall Definition 3, clause (i):

The situation s forces the proposition P within U_S iff for every $w \in U_S$ such that $s \subseteq w$ it holds that $w \in P$.

By this definition any basis s will (trivially) force p , if p is contained in the basis. Hence, deleting p from the basis will guarantee that p is no longer forced. So, the task is to reduce the basis s of $w_{@}$ ³² such that it no longer contains p , while to simultaneously conserve the rest of the $w_{@}$ -basis in order to ensure maximal similarity to $w_{@}$. After all, you want to maintain as much information about the actual state as possible when making a counterfactual assumption. This reduced

³²For simplicity I will refer to the set of possibly actual worlds as 'the' actual world $w_{@}$.

CHAPTER 3. THE FORMAL FRAMEWORK

basis (or, ‘pseudo basis’ as I will call it occasionally³³) of $w_{@}$ will then act as the kernel of the basis of any counterfactual world.³⁴ Updating state $\mathbf{1}[\Box(p \rightarrow q)][\Box(\neg p \rightarrow q \vee \neg q)][p]$ with the counterfactual assumption that $[\neg p]$ will hence involve three steps: in the first step the fact that p is eliminated from the basis of $w_{@}$. In the second step the counterfactual worlds are checked. More specifically, one has to identify the bases that maximally agree with the reduced basis of $w_{@}$ (or, to be more precise, of $w_{@} \downarrow p$ for that matter). And in the third step it remains to be checked whether for *all* these counterfactual worlds the claim put forward by the counterfactual is actually true. (For the particular example of Tichý 1 recall that the bases of $w_{@}$ are the sets $\{\langle p, 1 \rangle, \langle r, 0 \rangle\}$ for w_6 and $\{\langle p, 1 \rangle, \langle r, 1 \rangle\}$ for w_7 . From these the information that $\langle p, 1 \rangle$ will have to be retracted. Hence in the case under consideration the bases of the counterfactual worlds have to contain either $\langle r, 0 \rangle$ or $\langle r, 1 \rangle$. Having established this it is then easy to verify (when consulting the above given list of bases) that indeed *all* counterfactual worlds are at least partially based on either $\langle r, 0 \rangle$ or $\langle r, 1 \rangle$.)³⁵

The real deal: how things actually work

Sticking closer to Veltman’s account we will see that the three steps just described – though yielding the proper result – deliver only a rough heuristics. In Veltman’s theory the issue is solved with greater formal rigor, and it is necessary to elaborate more on the complexities of his account. The main difference between the above simplifying solution and Veltman’s principled account is that – unlike pretended above in the second step – the shift from the factual to the counterfactual is not

³³See footnote 17, p. 88 again for why the notion of a *basis* does not really apply here.

³⁴The point just made rests on the assumption that p figures as an independent premise. In case p derives from another premise p' , this other premise needs to be discarded from the basis in order to guarantee that p is not forced. In this case things become arguably more involved. Besides the nature of the nomological relation between p' and p that may play a role here also the underlying logic may play a decisive role.

³⁵Note again that r is merely stipulated for technical reasons. In the given scenario the proposition r is actually irrelevant. The process of reasoning from the bases is however facilitated by the assumption of an independent premise r . Otherwise the retraction of p would have resulted in the empty set \emptyset . While in principle feasible, this complicates things unnecessarily: it would have necessitated to consider *all* worlds in the further computation. This is alluded to in footnote 15 in Veltman [2005], p. 167, where Veltman remarks: “If there is no world $w \in U_S$ such that $s \subseteq w$, then, according to this definition [Definition 3 (i), D.B.], the situation s forces every proposition.”

CHAPTER 3. THE FORMAL FRAMEWORK

simply a magical shift in attention, but *the effect of an update*.

Since clause (i) of Definition 3 determines to consider a subset of the bases of $w_{@}$, the retraction of p in Tichý 1 includes not just – as suggested in the heuristics – the counterfactual worlds w_0, w_1, w_2, w_3 , but also those worlds that designate $w_{@}$, namely w_6 and w_7 . The corresponding retraction-state $(\mathbf{1}[\Box(p \rightarrow q)][\Box(\neg p \rightarrow q \vee \neg q)][p]) \downarrow \llbracket p \rrbracket$ is thus:

⟨	U_S	p	q	r	,	F_S	p	q	r	⟩
	w_0	0	0	0		w_0	0	0	0	
	w_1	0	0	1		w_1	0	0	1	
	w_2	0	1	0		w_2	0	1	0	
	w_3	0	1	1		w_3	0	1	1	
	w_4	/N	/Q	/Q		w_4	/N	/Q	/Q	
	w_5	/N	/Q	/N		w_5	/N	/Q	/N	
	w_6	1	1	0		w_6	1	1	0	
	w_7	1	1	1		w_7	1	1	1	

(Trivially, the worlds constituting this state are all nomologically sound and to a certain degree similar to the actual world $w_{@}$.)

The difference to the formulated heuristics is that only after an update with $\neg p$ (i.e. the assumption that the weather had not been bad, but fine) are the *bad weather*-worlds w_6 and w_7 eliminated. (And only then is a distinction between the factual and the counterfactual worlds achieved!) The state $((\mathbf{1}[\Box(p \rightarrow q)][\Box(\neg p \rightarrow q \vee \neg q)][p]) \downarrow \llbracket p \rrbracket) [\neg p]$ that is thus reached is the counterfactually updated state S [*If it had been the case that the weather was fine*]:

CHAPTER 3. THE FORMAL FRAMEWORK

U_S	p	q	r
w ₀	0	0	0
w ₁	0	0	1
w ₂	0	1	0
w ₃	0	1	1
w₄	/N	/Q	/Q
w₅	/N	/Q	/N
w ₆	1	1	0
w ₇	1	1	1

,

F_S	p	q	r
w ₀	0	0	0
w ₁	0	0	1
w ₂	0	1	0
w ₃	0	1	1
w₄	/N	/Q	/Q
w₅	/N	/Q	/N
w₆	/N	/N	/Q
w₇	/N	/N	/N

Having reached this state a position is gained that allows to finally assess the truth value of the counterfactual conditional. In order to do so it remains to be seen if the counterfactual's claim is legit and the state indeed supports the consequent of the counterfactual. In other words, one has to check whether *q* is *accepted* in the state, as Veltman puts it (cf. Veltman [2005], p. 160). Or, more formally:

$$((\mathbf{1}[\Box(p \rightarrow q)][\Box(\neg p \rightarrow q \vee \neg q)][p]) \downarrow [\Box][\neg p]) \models q$$

Recall that for the scenario of Tichý 1 the counterfactual *If the weather had been nice, Jones would have been wearing his hat* is intuitively unacceptable. It is not true. And, looking at the state depicted above, we see that indeed, unlike advanced by the counterfactual, *not all* of the worlds one has to consider are *Jones is wearing his hat*-worlds. While in worlds *w*₂ and *w*₃ the sentence is true, *q* turns out to be false in worlds *w*₀ and *w*₁. Hence, the counterfactual's universal claim – its determinacy with respect to the scenario it describes – *fails*. It does not cohere with what we know about the scenario. This in turn means that Veltman's account is capable to correctly reconstruct the initial intuition connected to Tichý's original example and the meaning of the counterfactual in question.

The idea of acceptance that is so crucial for the assessment of counterfactuals finds a different rendering when Veltman proposes to conceive of counterfactuals as *tests*. This is elaborated in Definition 5 ff.:

CHAPTER 3. THE FORMAL FRAMEWORK

Definition 5 (Counterfactuals as tests)

$S[\textit{if had been } \varphi, \textit{ would have been } \psi] = S$, if $S[\textit{if had been } \varphi] \models \psi$
 $S[\textit{if had been } \varphi, \textit{ would have been } \psi] = \mathbf{0}$, otherwise.

(Veltman (2005), p. 170)

3.4.2 Tichý 2

Before I will ultimately turn to the formal analysis of the data of Veltman's puzzle that provided the very starting point of this inquiry I will first complete the discussion of how Veltman's semantics works by applying it, as promised, also to Tichý 2. Again, I will proceed step-by-step, but this time I will not comment on every detail as extensively as in the case of Tichý 1. The major concerns will become clearly recognizable nonetheless. I will proceed as before: after giving a translation of the scenario of Tichý 2 into a propositional language I will run through the succession of updates and the computation of the bases. Finally the outcome of the enterprise is assessed. We will see that also in this case – which is more complex than Tichý 1 – the system is able to successfully reconstruct the intuitive valuation.

Translating the scenario

The main tenets of the story line of Tichý 2 translate to the below propositional atoms:

It is raining = p

Jones is wearing his hat = q

The coin comes up heads = r

Adhering to Veltman's formalization (cf. Veltman (2005), p. 170), the nomological core of the scenario is captured in one boxed formula :

Jones is wearing his hat if and only if the weather is bad or the coin comes up heads = $\Box(q \leftrightarrow (p \vee r))$

CHAPTER 3. THE FORMAL FRAMEWORK

The immediately given facts particular to the scenario are:

The coin comes up heads = $\langle r, 1 \rangle$

It is raining = $\langle p, 1 \rangle$

Given the nomological and factual premises, the following information is derivable for the present situation:

Jones is wearing his hat = $\langle q, 1 \rangle$

The succession of nomological and factual updates

As before, the update process shall depart from the minimal state **1**.³⁶

⟨	U_S	p	q	r	,	F_S	p	q	r	⟩
	<i>w</i> ₀	0	0	0		<i>w</i> ₀	0	0	0	
	<i>w</i> ₁	0	0	1		<i>w</i> ₁	0	0	1	
	<i>w</i> ₂	0	1	0		<i>w</i> ₂	0	1	0	
	<i>w</i> ₃	0	1	1		<i>w</i> ₃	0	1	1	
	<i>w</i> ₄	1	0	0		<i>w</i> ₄	1	0	0	
	<i>w</i> ₅	1	0	1		<i>w</i> ₅	1	0	1	
	<i>w</i> ₆	1	1	0		<i>w</i> ₆	1	1	0	
	<i>w</i> ₇	1	1	1		<i>w</i> ₇	1	1	1	

Updating the minimal state with the rule that *Jones is wearing his hat if and only if the weather is bad or the coin comes up heads* yields state **1** [$\Box(q \leftrightarrow (p \vee r))$].

Those worlds that violate the law are discarded from the logical space:

³⁶I will refrain from assuming (as has been done before) an additional, independent fact, since the ‘danger’ of having to deal with empty pseudo-bases will not arise, as we will see in a moment. Things would become unnecessarily complicated without adding any further insight.

CHAPTER 3. THE FORMAL FRAMEWORK

U_S	p	q	r
<i>w</i> ₀	0	0	0
<u><i>w</i>₁</u>	<u>/0/</u>	<u>/0/</u>	<u>/1/</u>
<u><i>w</i>₂</u>	<u>/0/</u>	<u>/1/</u>	<u>/0/</u>
<i>w</i>₃	0	1	1
<u><i>w</i>₄</u>	<u>/1/</u>	<u>/0/</u>	<u>/0/</u>
<u><i>w</i>₅</u>	<u>/1/</u>	<u>/0/</u>	<u>/1/</u>
<i>w</i> ₆	1	1	0
<i>w</i> ₇	1	1	1

F_S	p	q	r
<i>w</i> ₀	0	0	0
<u><i>w</i>₁</u>	<u>/0/</u>	<u>/0/</u>	<u>/1/</u>
<u><i>w</i>₂</u>	<u>/0/</u>	<u>/1/</u>	<u>/0/</u>
<i>w</i>₃	0	1	1
<u><i>w</i>₄</u>	<u>/1/</u>	<u>/0/</u>	<u>/0/</u>
<u><i>w</i>₅</u>	<u>/1/</u>	<u>/0/</u>	<u>/1/</u>
<i>w</i> ₆	1	1	0
<i>w</i> ₇	1	1	1

Providing the information that indeed heads came up (= $\langle r, 1 \rangle$) results in a change in the factual dimension of the state: it is narrowed down to worlds w_3 and w_7 (highlighted with **bold** font) while worlds w_0 and w_6 turn out to be counterfactual (underlined and in **red**). The result $\mathbf{1}[\Box(q \leftrightarrow (p \vee r))][r]$ is as follows:

U_S	p	q	r
<i>w</i> ₀	0	0	0
<u><i>w</i>₁</u>	<u>/0/</u>	<u>/0/</u>	<u>/1/</u>
<u><i>w</i>₂</u>	<u>/0/</u>	<u>/1/</u>	<u>/0/</u>
<i>w</i>₃	0	1	1
<u><i>w</i>₄</u>	<u>/1/</u>	<u>/0/</u>	<u>/0/</u>
<u><i>w</i>₅</u>	<u>/1/</u>	<u>/0/</u>	<u>/1/</u>
<i>w</i> ₆	1	1	0
<i>w</i> ₇	1	1	1

F_S	p	q	r
<u><i>w</i>₀</u>	<u><u>0</u></u>	<u><u>0</u></u>	<u><u>0</u></u>
<u><i>w</i>₁</u>	<u>/0/</u>	<u>/0/</u>	<u>/1/</u>
<u><i>w</i>₂</u>	<u>/0/</u>	<u>/1/</u>	<u>/0/</u>
<i>w</i>₃	0	1	1
<u><i>w</i>₄</u>	<u>/1/</u>	<u>/0/</u>	<u>/0/</u>
<u><i>w</i>₅</u>	<u>/1/</u>	<u>/0/</u>	<u>/1/</u>
<u><i>w</i>₆</u>	<u><u>1</u></u>	<u><u>1</u></u>	<u><u>0</u></u>
<i>w</i>₇	1	1	1

The next update is due to the information that *It is raining* being communicated. This further reduces the set of the possibly actual worlds from $\{w_3, w_7\}$ to $\{w_7\}$, since w_3 is a world in which the statement *It is raining* is false. Hence the state $\mathbf{1}[\Box(q \leftrightarrow (p \vee r))][r][p]$ is generated:

CHAPTER 3. THE FORMAL FRAMEWORK

⟨	U_S	p	q	r	, ⟨	F_S	p	q	r
	w_0	0	0	0		w_0	<u>0</u>	<u>0</u>	<u>0</u>
	w_1	/0	/0	/1		w_1	/0	/0	/1
	w_2	/0	/1	/0		w_2	/0	/1	/0
	w_3	0	1	1		w_3	<u>0</u>	<u>1</u>	<u>1</u>
	w_4	/1	/0	/0		w_4	/1	/0	/0
	w_5	/1	/0	/1		w_5	/1	/0	/1
	w_6	1	1	0		w_6	<u>1</u>	<u>1</u>	<u>0</u>
	w_7	1	1	1		w_7	1	1	1

Computing the bases

Again, before finally addressing the counterfactual update first the bases of the factual and counterfactual worlds that have been established so far shall be specified. We will see that the line of reasoning instrumental in the identification of the bases crucially involves the interplay of the bi-conditional and of the disjunction that occur in the law $\Box(q \leftrightarrow (p \vee r))$ that governs the scenario's horizon.

The basis of w_7 is $\{\langle p, 1 \rangle, \langle r, 1 \rangle\}$. From this it is deducible – by the rule $\Box(q \leftrightarrow (p \vee r))$ being true and the definitions of ‘ \vee ’ and ‘ \leftrightarrow ’ – that $\langle q, 1 \rangle$.³⁷ The line of reasoning is as follows:

0. $\langle r, 1 \rangle$
1. $\langle p, 1 \rangle$
2. $\langle \Box(q \leftrightarrow (p \vee r)), 1 \rangle$
3. $\langle (p \vee r), 1 \rangle$ by 0. (and 1.) and definition of ‘ \vee ’
4. $\langle q, 1 \rangle$ by 3. and 2. and definition of ‘ \leftrightarrow ’

The basis of w_6 is $\{\langle p, 1 \rangle, \langle r, 0 \rangle\}$. The world w_6 is determined by and large as above, with the only difference being the value given for r .³⁸

³⁷Actually, since p and r are given, each of both facts alone suffices to derive the value of q . Note however that, since neither fact alone allows to determine the other facts value, p and r are not fully independent. The basis of the world is thus arguably given only by both, p and r .

³⁸Again the atoms p and r are not mutually independent.

CHAPTER 3. THE FORMAL FRAMEWORK

0. $\langle r, 0 \rangle$
1. $\langle p, 1 \rangle$
2. $\langle \Box(q \leftrightarrow (p \vee r)), 1 \rangle$
3. $\langle (p \vee r), 1 \rangle$ by 1. (and 0.) and definition of ' \vee '
4. $\langle q, 1 \rangle$ by 3. and 2. and definition of ' \leftrightarrow '

Or, alternatively, $\{\langle q, 1 \rangle, \langle r, 0 \rangle\}$:

0. $\langle r, 0 \rangle$
1. $\langle q, 1 \rangle$
2. $\langle \Box(q \leftrightarrow (p \vee r)), 1 \rangle$
3. $\langle (p \vee r), 1 \rangle$ by 1. and 2. and definition of ' \leftrightarrow '
4. $\langle p, 1 \rangle$ by 3. and 0. and definition of ' \vee '

Depending on the 'direction' of the possible derivations w_0 has two bases: when departing from the right side of $\Box(q \leftrightarrow (p \vee r))$ the basis is $\{\langle p, 0 \rangle, \langle r, 0 \rangle\}$; when starting on the left side, the basis is $\{\langle q, 0 \rangle\}$. First the right side derivation is given:³⁹

0. $\langle r, 0 \rangle$
1. $\langle p, 0 \rangle$
2. $\langle \Box(q \leftrightarrow (p \vee r)), 1 \rangle$
3. $\langle q, 0 \rangle$ by 0. and 1. and definition of ' \vee ', and 2. and definition of ' \leftrightarrow '

The alternative basis $\{\langle q, 0 \rangle\}$ derives the values of p and r as follows:

0. $\langle q, 0 \rangle$
1. $\langle \Box(q \leftrightarrow (p \vee r)), 1 \rangle$
2. $\langle (p \vee r), 0 \rangle$ by 0. and 1. and definition of ' \leftrightarrow '
3. $\langle p, 0 \rangle$ and $\langle r, 0 \rangle$ by 2. and definition of ' \vee '

Two bases for w_3 are $\{\langle p, 0 \rangle, \langle q, 1 \rangle\}$ and $\{\langle p, 0 \rangle, \langle r, 1 \rangle\}$. The former derives the value of p :

³⁹In case more than one basis exists the question whether one is to be favored over another arises of course. This point however can not be addressed here.

CHAPTER 3. THE FORMAL FRAMEWORK

0. $\langle p, 0 \rangle$
1. $\langle q, 1 \rangle$
2. $\langle \Box(q \leftrightarrow (p \vee r)), 1 \rangle$
3. $\langle p \vee r, 1 \rangle$ by 1. and 2. and definition of ' \leftrightarrow '
3. $\langle r, 1 \rangle$ by 3. and 0. and definition of ' \vee '

The second basis of w_3 allows to determine the truth of q :

0. $\langle p, 0 \rangle$
1. $\langle r, 1 \rangle$
2. $\langle \Box(q \leftrightarrow (p \vee r)), 1 \rangle$
3. $\langle p \vee r, 1 \rangle$ by 0. and 2. and definition of ' \vee '
4. $\langle q, 1 \rangle$ by 2. and 3. and definition of ' \leftrightarrow '

The counterfactual update

Returning to the update process after the digression concerned with the specification of the bases, the counterfactual update takes center stage. Recall that after completing the nomological and factual updates in $w_{@}$ it is the case that Jones is wearing his hat, because the coin came up heads and it is raining.

Since the counterfactual *If the weather had been fine, Jones would have been wearing his hat* demands to consider those worlds in which the weather is fine, the counterfactual worlds to look at are those in which $\langle p, 0 \rangle$ holds, instead of the actual fact $\langle p, 1 \rangle$. (Again, one has to assume a strong antonymous relation between the weather being fine and the weather being bad in the sense that the sentence *It is raining* being false implies *The weather is fine*, or, even better yet, *The sun is shining*.)

When implementing Veltman's proposal to the analysis of the scenario that the counterfactual demands to consider the first step involved is to retract $\langle p, 1 \rangle$ from the basis of $w_{@}$. Since w_7 is the only world left in the set of the possibly actual worlds, its basis is thus to be altered accordingly. Hence, departing from $\{\langle p, 1 \rangle, \langle r, 1 \rangle\}$ the kernel of the counterfactual worlds to consider is given by the

CHAPTER 3. THE FORMAL FRAMEWORK

assignment $\langle r, 1 \rangle$.

The *retraction*-state $(\mathbf{1}[\Box(q \leftrightarrow (p \vee r))][r][p])\downarrow \llbracket p \rrbracket$ thus collects every world in which $\langle r, 1 \rangle$ is contained in a maximal subset of its basis that does not force p to be true. The only worlds that satisfy this condition are w_3 and w_7 . Hence, the *retraction*-state is:

⟨	U_S	p	q	r	,	F_S	p	q	r	⟩
	w_0	0	0	0		w_0	0	0	0	
	w_1	0	0	1		w_1	0	0	1	
	w_2	0	1	0		w_2	0	1	0	
	w_3	0	1	1		w_3	<u>0</u>	<u>1</u>	<u>1</u>	
	w_4	1	0	0		w_4	1	0	0	
	w_5	1	0	1		w_5	1	0	1	
	w_6	1	1	0		w_6	1	1	0	
	w_7	1	1	1		w_7	<u>1</u>	<u>1</u>	<u>1</u>	

Next, the counterfactual assumption that *not p* has to be added to this state. –Which in turn leads to the suspension of the actual world w_7 . Doing so one hence achieves state $((\mathbf{1}[\Box(q \leftrightarrow (p \vee r))][r][p])\downarrow \llbracket p \rrbracket)[\neg p]$:

⟨	U_S	p	q	r	,	F_S	p	q	r	⟩
	w_0	0	0	0		w_0	0	0	0	
	w_1	0	0	1		w_1	0	0	1	
	w_2	0	1	0		w_2	0	1	0	
	w_3	0	1	1		w_3	<u>0</u>	<u>1</u>	<u>1</u>	
	w_4	1	0	0		w_4	1	0	0	
	w_5	1	0	1		w_5	1	0	1	
	w_6	1	1	0		w_6	1	1	0	
	w_7	1	1	1		w_7	1	1	1	

While this completes what it means – according to Veltman – to make a counterfactual assumption, it remains to be checked whether this assumption also supports $\langle q, 1 \rangle$. Only if this is the case for w_3 will the counterfactual be acceptable. Since it is indeed the case that in w_3 it holds that Jones is wearing his hat, the counterfactual may be considered to be true. Hence:

CHAPTER 3. THE FORMAL FRAMEWORK

$$((\mathbf{1}[\Box(q \leftrightarrow (p \vee r))][r][p])\downarrow \llbracket p \rrbracket)[\neg p] \models q$$

And this exactly coincides with the intuitive interpretation of the counterfactual in the given scenario. Veltman's theory is thus also in this case able to correctly reconstruct the intuition connected to the example.

3.5 Summary

Having completed the introduction of the formal framework in which the pending analysis is to be formulated, those ideas and concepts of Veltman's theory that will be of particular importance for the account to be developed shall be summarized now. The actual concern and result of Veltman, namely his theory of what it means to make counterfactual assumptions, will be only of secondary importance in this. At least insofar as nothing will be added or revised in this account. The semantics of the implicit antecedents that occurs in the discourses of Veltman's puzzle will be accounted for – strictly adhering to Veltman – in terms of how he conceives what it means to make counterfactual assumptions, namely involving the achievement of the retraction state and the execution of the counterfactual update. The motivation and concepts that underlie these steps shall however be quickly recalled.

First to be mentioned is the idea to conceive of counterfactual conditionals as *support tests* which emulates the simple intuition that a counterfactual is true whenever in the state reached by making the counterfactual assumption also the conclusion expressed in the consequent is true. What is actually tested is whether the assumption represents a tenable hypothesis. It is so, whenever the conclusion that it derives is consistent with the informational background of the assumption. The notion of the support test may however also be considered to be relevant with respect to the peculiar role that counterfactuals play in conversation. Agreeing upon the truth of a counterfactual indicates not only the consistency of an assumption with a state, but that moreover the individual information states of the participants of communication are indeed aligned. Hence, one may indeed assume that coun-

CHAPTER 3. THE FORMAL FRAMEWORK

terfactuals serve to test the common ground.

The second point to be mentioned is already inherent to the first point and actually motivates the technicalities of Veltman's theory. (Foremost the definition of the retraction state and the concept of the base.) The discussion of the Tichý-examples highlighted that making a counterfactual assumption involves not only the revision of bare facts, but that it is decisive for the truth of the counterfactual that in the assumption also relations between facts are recognized. As already mentioned above, in the formulation and assessment of a counterfactual the (nomological) background is of utmost importance. We will see later on in the analysis that in cases that completely lack such a nomological backbone a heavy dose of pragmatic reasoning is necessary in order to be able to consider in such a case the communication of a counterfactual to be rational at all. At this point the analysis will indeed go beyond giving an account of the felicity conditions of the elliptic discourses of Veltman's puzzle that acted as the original starting point. In addition it will also provide with a deeper look at certain issues in the pragmatics and semantics of counterfactuals that goes beyond what is laid out by Veltman (and, as far as I can see, the other literature on counterfactuals).

Chapter 4

The Analysis

Turning now to the analysis the groundwork that has been laid down in the preceding three chapters shall quickly be recalled. The data under concern and the problems they pose were introduced in the first chapter. The basis gained there was then located within its wider empirical context in the second chapter. And in the third chapter the formalism was expounded which will now be employed in the analysis of the issues formulated in chapter 1. The starting point for the pending formal account will be provided by one of the very first examples drawn from Veltman (2005). It is the familiar discourse

(1) John did not drink any wine. He would have become sick.

While in due course also its infelicitous counterpart

(2) # John drank too much wine. He would not have become sick.

will be accounted for, the chapter has more to offer than just an analysis of these two examples. Its structure is summarized in the following outline.

Outline of the chapter

First, the procedural details of the interpretation of the felicitous example (1) will be spelled out in depth in section 4.1 by computing the sequence of updates required. The example represents a paradigmatic case for the successful processing

CHAPTER 4. THE ANALYSIS

of a felicitous discourse. Next, I will turn in section 4.2 to the actual semantic analysis. Here I will not just consider example (1), but will also attend to its infelicitous counterpart, example (2), and a number of variants of both. These include for instance the felicitous version of (2) with an explicit antecedent, and a version of (1) in which the presuppositions of the scenario are canceled. The discussion of both of these cases will provide with important insights regarding the semantic and pragmatic resources exploited in the interpretation of counterfactuals in general. The next major point of attention is addressed in section 4.3 in which the semantic analysis of example (2) is complemented with a clarification of the role that a conjunctive vs. a disjunctive premise set plays in general for the felicity of the discourses. This continues the discussion of the idea that an exhaustivity constraint might be crucial for the analysis of the data with which the very first chapter of this thesis was completed. Finally, at the end of the analysis, in section 4.4 the availability of a *because*-paraphrase for most of the examples is studied in detail.

4.1 The procedural side of discourse computation

As already mentioned at the start of the analysis I will return to the original example *John did not drink any wine. He would have become sick.*¹

From what has been discussed so far in the preceding chapters, it ought to be clear that things are not as simple, as the surface structure of the discourse suggests. Speaking more formally, this sequence of two sentences (or, when acknowledging the special nature of the *would have* + Past Participle-complex, the sequence of a sentence and a sentence-like expression), can not be understood as a sequence of two simple updates, with the initial sentence acting as the first update, and the subsequent counterfactual relict-clause as the second update.

Not only is the example empirically more complex than it appears at first sight in that it is elliptic and involves a negation-induced alternative that acts as a contextu-

¹For reasons of brevity I will in the following reduce the discourse to *John did not drink. He would have become sick.*

CHAPTER 4. THE ANALYSIS

ally determined implicit antecedent which completes a counterfactual conditional, but also theoretically is it rather demanding and necessitates the application of the full update machinery for counterfactuals.

Indeed, following Veltman's account, the discourse comprises a whole sequence of updates that eventually cumulates in a support test. This history of semantic composition and processing is neatly summarized in the formal representation of the state designating the support test to be achieved. This state gives the terminal point of the interpretation process. It can be given as:

discourse context $[John\ did\ not\ drink] \downarrow [John\ did\ not\ drink] [John\ drank]_{=NinA}$
 $\models John\ would\ have\ become\ sick$

The procedural steps that lead up to this state will now be discussed in detail. The starting point is provided by the assumption that the cognitive agent is in the minimal state.

4.1.1 The communicative environment of the discourse in the initial state

The computation of the discourse's meaning shall depart from the minimal state:

$$\mathbf{1} = \langle W, W \rangle$$

Nomological updates transform the minimal state to a state $\mathbf{1}^+$:

$$\mathbf{1}^+ = \langle U_S, F_S \rangle$$

This state $\mathbf{1}^+$ is distinguished from the minimal state by the assumption that the state's cognitive bearer established at least some laws. The exact nature of this horizon shall however remain irrelevant. (With the exception of the question whether the content of the counterfactual that is implicit in the discourse under discussion, i.e. the empirical truth that *Drinking too much alcohol will result in sickness*, can be presupposed or not. This point will be addressed separately later on in the section on radical interpretation of counterfactuals on page 123 ff.) What

CHAPTER 4. THE ANALYSIS

is noteworthy with respect to $\mathbf{1}^+$ is that at this stage F_S is still identical to U_S ($F_S = U_S$), since so far no other facts than laws have been established.

Although it is redundant (and admittedly not quite realistic) I will assume that factual updates separately follow the nomological updates. These factual updates yield a state $\mathbf{1}^\oplus$:

$$\mathbf{1}^\oplus = \langle U_S, F_S \rangle$$

Unlike in state $\mathbf{1}^+$ in state $\mathbf{1}^\oplus$ the set of facts F_S represents now a proper subset of U_S : $F_S \subset U_S$. This state provides the nomological and factual environment into which the discourse under discussion will have to be incorporated.

4.1.2 Updating with a NinA-trigger: the discourse initial sentence *John did not drink*

Having thus defined the communicative context in which the discourse is to be embedded, next the update of state $\mathbf{1}^\oplus$ with the discourse initial sentence shall be considered. The concrete situation is thus that of some agent A engaging in communication with another individual B by uttering *John did not drink*. Sticking to Definition 1 (i) in Veltman (2005) on p. 165 and Veltman's subsequent elaboration the basic meanings involved are:

$$\llbracket \text{John drank} \rrbracket = \{w \in W \mid w(\text{John drank}) = 1\}^2$$

$$\llbracket \text{not } \varphi \rrbracket = W \setminus \llbracket \varphi \rrbracket^3$$

Based on this, Veltman's notion of *interpretation* (as specified in Veltman (2005), p. 167, Definition 2) accounts for the actual effect of updating state $\mathbf{1}^\oplus$ with the proposition *John did not drink*. The resulting state is:

$$\mathbf{1}^\oplus[\text{John did not drink}] = \langle U_S, F_S \cap \llbracket \text{John did not drink} \rrbracket \rangle$$

²The more standard alternative is of course: $\llbracket \text{John drank} \rrbracket = \{w \in W \mid \text{John drank}(w) = 1\}$. As has already been mentioned in my footnote 11 in section 3.3.2 on page 78 this definition impedes a straightforward translation into a standard type-theoretical semantics.

³Note that as long as only factual information is negated (and not nomological propositions!) one might restrict the domain of negation to F_S , and not W as a whole.

CHAPTER 4. THE ANALYSIS

As with any update also this update ultimately results in a partition of F_S that differentiates the factual from the counterfactual dimension. The counterfactual half of F_S is complementary to the factual dimension and trivially consists of all worlds within U_S that are – just like the NinA-worlds – not factual. Shifting the perspective to what is (known to be) not the case (indicated by the superscript \ominus), the counterfactual dimension of the cognitive state of the agent can be given as follows:

$$\mathbf{1}^{\ominus}[\textit{John did not drink}] = \langle U_S, U_S \setminus (F_S \cap \llbracket \textit{John did not drink} \rrbracket) \rangle$$

Considering the dichotomous nature of any update, strictly speaking both, $\mathbf{1}^{\oplus}[\textit{John did not drink}]$ and $\mathbf{1}^{\ominus}[\textit{John did not drink}]$, simultaneously represent the cognitive state that some individual B achieves after another individual A communicates the sentence *John did not drink* to B and B successfully interprets this information.⁴

4.1.3 Communicating *He would have become sick*

Next to the statement *John did not drink* follows A commenting to B *He would have become sick*.⁵ The complex of *would have* + Past Participle marks two things: first, that *He would have become sick* does not represent a factual update, and second – following Veltman’s conception and his terminology –, that it rather acts as a *support test* (cf. Veltman (2005), p. 170ff.) on the counterfactual domain (understanding that the complex figures as the overt consequent of an implicit counterfactual conditional). The claim that A connects to this part of the discourse can be understood as A ascertaining that when s/he adds the counterfactual assumption that *John had drunk something* to what s/he already knows about the world, then it logically follows from this that *John would have become sick*. I.e. A ’s state of information supports the conclusion that *John would have become sick*, if it is assumed that *John drank*.

⁴The question whether the minimal state is also of a dichotomous nature seems to be rather intriguing, but shall not be addressed here.

⁵In the following I assume that the anaphoric pronoun *he* has been successfully resolved for its antecedent *John*.

CHAPTER 4. THE ANALYSIS

Elaborating on the communicative function of the discourse allows to give the notion of *support* additional intuitive appeal. The function is either – in case *B* agrees on the validity of the support test – the mutual reassurance of a shared common ground of the participants of communication, or – in case *B* disagrees on the test’s validity, e.g. by disputing the truth of the consequent – to establish the defectiveness of the presupposed common ground in order to facilitate the construction of a more congruent notion of a shared background by discussing the point further.⁶

In order for *B* to successfully execute the support test implicitly proposed by *A*, the following prerequisites are necessary:

- A suitable context is required that facilitates the interpretation of the counterfactual consequent *He would have become sick*. This context is the agent’s cognitive state *S*.

and

- For this context it must hold that it supports *He (= John) would have become sick*:

$$S \models \llbracket He_{(=John)} \text{ would have become sick} \rrbracket$$

Or, more accurately, merely recognizing the propositional content of the *would have* + Past Participle complex:

$$S \models \llbracket He_{(=John)} \text{ is sick} \rrbracket$$

(Although the distinction is debatable, one might argue that the first point exploits the formal-grammatical properties of the *would have* + Past Participle construction in terms of its specific mood, while the second point takes into account the

⁶Veltman comments on this issue: “By asserting ‘*if had been φ , would have been ψ* ’, a speaker makes a kind of a comment: ‘Given the general laws and the facts I am acquainted with, the sentence ψ is supported by the state I get in when I assume that φ had been the case’. The addressee is supposed to determine whether the same holds on account of his or her own information. If not, a discussion will arise, and in the course of this discussion both the speaker and the hearer may learn some new laws and facts, which could affect the outcome of the test.” (Veltman (2005), p. 171)

CHAPTER 4. THE ANALYSIS

peculiar semantics deriving from the complex.)

The question is thus whether there is (with respect to B) a context (read: cognitive state) that is accessible and suitable for the interpretation of the counterfactual consequent He_{John} *would have become sick*?

4.1.4 Interpreting *He would have become sick* – a naive first attempt

The answer to this question is provided by the discourse itself. The complementary correlate to state $\mathbf{1}^\ominus[John\ did\ not\ drink]$, i.e. the state's counterfactual dimension given with $\mathbf{1}^\ominus[John\ did\ not\ drink]$ represents a context that formally meets the requirements specified. It is a grammatically suitable context because it is counterfactual, thus satisfying the constraint set forth by mood, and it is logically suitable because it is specific enough to actually conduct the support test in a non-fuzzy way (since it is distinguished by the saliency of the NinA producing update). It thus provides the basis for solid inference.

Put tentatively and very naively (from the discussion in the preceding chapter it should be clear that the perspective to be taken is not completely accurate) the discourse puts forth that the following support relation is to be entertained by the interpreter:

$$\langle U_S, U_S \setminus (F_S \cap \llbracket John\ did\ not\ drink \rrbracket) \rangle \models He\ became\ sick$$

Slightly shifting the perspective, this sums up to the expectation (of A) that after updating the irreal background with the counterfactual assumption that *John drank*, B agrees on the truth of the proposition *John would have become sick*:

$$background_{irreal} \cap John\ drank_{irreal} \models He\ became\ sick_{irreal}$$

Note that if one highlights the role and the origin of the implicit antecedent, this is equivalent to:

$$background_{irreal} \cap NinA(discourse\ initial\ sentence) \models He\ became\ sick_{irreal}$$

CHAPTER 4. THE ANALYSIS

Based on this, it suggests itself to represent the meaning of the subjunctive in form of the *would have* + Past Participle complex very roughly as follows:⁷

$$\llbracket \textit{would have} + \textit{PastParticiple} \rrbracket = \lambda p \lambda H. H \models p$$

In this the variable p stands for a proposition, while H gives a context (or background, or cognitive state) in terms of a set of propositions. This crude ‘semantics’ for the subjunctive seeks to postulate that whatever context it is that the proposition p (that is marked by the subjunctive) is asserted in, it is such that it stands in the *support relation* to p . What is important to note is that the involved context or background is required to be *counterfactual*, i.e. it holds $H = S = \langle U_S, U_S \setminus F_S \rangle$ (or, more generally, $W \setminus F_S$ ⁸). (Note that this conception of a context as a set of propositions is in full accordance with the definitions of the notions of *support* resp. *logical validity* given in Veltman (2005), p. 160.)

4.1.5 The revised full account

Conceived like this, the general constraints set forth above on the meaning and interpretation of the counterfactual consequent and its admissible contexts are reflected. However, the account is still far too crude to deliver satisfactory results. As already mentioned, it needs to be refined. In order to validate the support relation that is emulated in $\textit{background}_{\textit{irreal}} \cap \textit{NinA}(\textit{discourse initial sentence}) \models \textit{He became sick}_{\textit{irreal}}$ above, it does not suffice to simply consult ‘the counterfactual half of discourse’, as suggested by specifying the supporting context to be $\langle U_S, U_S \setminus (F_S \cap \llbracket \textit{John did not drink} \rrbracket) \rangle$. Indeed, the subjunctive does not refer indiscriminately to the counterfactual dimension. It is semantically way more subtle, selective and differentiate than this. Recall what the discussion of the Tichý-examples highlighted: also in case of counterfactual deliberation the premises of a scenario have to be acknowledged, if one is to reach valid, i.e. truth preserving conclusions. This means that not *any* counterfactual world will do, but only

⁷The aim is to give a provisional illustration of the idea. The representation actually is formally inaccurate and the λ -term undefined, since the use of ‘ \models ’ in the λ -term is a blatant mix of formal and meta-theoretical vernacular. In case Veltman’s logic had been supplied with a deduction theorem, a suitable formulation would be $\llbracket \textit{would have} + \textit{PastParticiple} \rrbracket = \lambda p \lambda H. H \rightarrow p$.

⁸Since (most of the time) one needs to be concerned with law-abiding worlds solely, $U_S \setminus F_S$ sufficiently designates the counterfactual dimension.

CHAPTER 4. THE ANALYSIS

those counterfactual worlds must be considered that are nomologically sound and cohere not just with the laws and rules, but also the bare facts constitutive of the situation. In other words, within the counterfactual dimension only those worlds are to be regarded that are maximally similar to the actual world. –Which is after all the world that acts as the point of reference for the counterfactual to begin with.

This means for the approach within Veltman’s framework that after having reached state $\mathbf{1}^{\oplus}[\textit{John did not drink}]$ and facing the subsequent comment *He would have become sick* one first has to compute the retraction state (i.e. the state that models those worlds that share the same maximal set of essential and indispensable premises as the actual world minus the fact to be countered):

$$\mathbf{1}^{\oplus}[\textit{John did not drink}] \downarrow [\textit{John did not drink}]$$

Following this, the retraction state next has to be updated with the counterfactual assumption that *John drank*, i.e. with the NinA that is generated by the discourse initial negative sentence *John did not drink*. The state reached is thus:

$$\mathbf{1}^{\oplus}[\textit{John did not drink}] \downarrow [\textit{John did not drink}] [\textit{John drank}]$$

Only after this update is completed is the state achieved for which it is supposed to hold that it supports the counterfactual consequent *He would have become sick* to the effect that it is accepted to be true in this context:

$$\mathbf{1}^{\oplus}[\textit{John did not drink}] \downarrow [\textit{John did not drink}] [\textit{John drank}] \models \textit{He would have become sick}$$

Paying heed to the intricacies of the semantics of counterfactuals, it becomes clear that the background before which the counterfactual consequent is interpreted is a grammatically required manipulated context, and not the crude counterfactual context as proposed above. It is the result of retraction and counterfactual update at the underlying bases.

However, the meaning given for *would have* + Past Participle above is in principle tenable, provided it is taken into account that the irreal background *H* specified therein is the result of proper counterfactual deliberation.

4.2 The Semantic account

4.2.1 The first case: elliptic discourse with nomological presupposition

To complete the picture it remains to be seen in detail what semantic effects the procedural steps described so far evoke on the logical space. This will be fleshed out now. First the example *John did not drink (any wine). He would have become sick* will be computed.

Unlike as before I will not present states and their updates in form of (a sequence of) ordered pairs. Instead I will merge the nomological and factual dimension of the state in a single table that also captures the succession of updates. The presentation in a single matrix is more compact, and avoids certain redundancies inherent to the account used so far while keeping tabs of everything. In its informational content it is equivalent to the notation using ordered pairs.

The table below models the computation of the discourse. Each column gives the result of the update of the preceding column to the left with the proposition given in the header (with the exception of the column headed *basis*, which specifies the basis of every world). The minimal state is provided by the truth value assignments for the propositions p , q , and r , with p translating the proposition *John drinks*, while q formalizes *John is sick*. The third variable r translates an arbitrary proposition that completes the scenario. I will assume that the law that *If you drink too much, you will become sick* is already established, i.e. can be presupposed as empirically confirmed. Hence, the minimal state $\mathbf{1}$ has already been updated with the information $\Box(p \rightarrow q)$. This allows to immediately discard worlds w_4 and w_5 from any further consideration. Hence, these worlds are marked ‘—’ in the table. Those worlds that remain within the logical space of the scenario are checked in the table. The resulting state $\mathbf{1}[\Box(p \rightarrow q)]$ represents the context into which the discourse initial sentence *John did not drink* will have to be embedded.

CHAPTER 4. THE ANALYSIS

	p	q	r	$\Box(p \rightarrow q)$	$\neg p$	basis	$\downarrow \llbracket \neg p \rrbracket$	p	$\models q$
w_0	0	0	0	\checkmark	fact	$\langle q, 0 \rangle, \langle r, 0 \rangle$	\checkmark	—	—
w_1	0	0	1	\checkmark	fact	$\langle q, 0 \rangle, \langle r, 1 \rangle$	\checkmark	—	—
w_2	0	1	0	\checkmark	fact	$\langle p, 0 \rangle, \langle q, 1 \rangle, \langle r, 0 \rangle$	\checkmark	—	—
w_3	0	1	1	\checkmark	fact	$\langle p, 0 \rangle, \langle q, 1 \rangle, \langle r, 1 \rangle$	\checkmark	—	—
w_4	1	0	0	—	—	—	—	—	—
w_5	1	0	1	—	—	—	—	—	—
w_6	1	1	0	\checkmark	ctfct.	$\langle p, 1 \rangle \langle r, 0 \rangle$	\checkmark	1	1
w_7	1	1	1	\checkmark	ctfct.	$\langle p, 1 \rangle \langle r, 1 \rangle$	\checkmark	1	1

Updating with the proposition *John did not drink* expressed by the discourse initial sentence distinguishes the factual from the counterfactual worlds (in the table abbreviated *fact* and *ctfct.* respectively). For each of these worlds is the basis specified in the next column. In the retraction state (given in the column $\downarrow \llbracket \neg p \rrbracket$) those premises that force the information that *John did not drink* are abandoned from the basis of the possibly actual worlds. Those worlds within U_S that share the remaining premises define the retraction state, i.e. the state to be updated with the counterfactual assumption. This state includes (trivially) the factual worlds, but also every counterfactual world that contains the identified ‘innocuous’ premises. These are all worlds except w_4 and w_5 , which, since they are impossible worlds, do not have to be considered anyway.

Next, the update with the negation induced alternative, i.e. the irreal proposition *John drank (wine)* is executed. Being effectively equivalent to making a counterfactual assumption, it allows to cancel the factual worlds and thus re-establishes the distinction between factual and counterfactual worlds. (The worlds that are canceled – this time the factual worlds – are again marked ‘—’ in the table.) Having achieved this state what is left to be seen is whether the remaining counterfactual worlds indeed support the counterfactual consequent *He_{John} would have become sick*. When consulting the remaining worlds w_6 and w_7 for the truth value assignments for p (i.e. *John drank*) and q (i.e. *John is sick*) (which are for immediacy repeated in the respective cells), it turns out that both are true in these worlds. Hence, the former in fact supports the latter (which should not come as a surprise,

CHAPTER 4. THE ANALYSIS

since given the law $\Box(p \rightarrow q)$ and the assumption that p , it follows by *Modus Ponens* that q). Updating the described state S with the counterfactual thus eventually only confirms the given state: $S[\text{If John had drunk any wine, he would have become sick}] = S$. For the counterfactual conditional *If John had drunk any wine, he would have become sick* this in turn means that under the given circumstances it is indeed true.

4.2.2 The second case: radical interpretation of counterfactuals – non-elliptic discourse without nomological presuppositions

Before analyzing the infelicitous discourse, first a variant of the felicitous example shall be considered that will shed some light on some very basic interpretive principles connected to counterfactuals. In this variant no law that relates drinking and sickness is presupposed and interpretation of the conditional *If John had drunk, he would have become sick* will become *radical* insofar the logical space into which the counterfactual is to be embedded is virtually unrestricted. The question how counterfactuals are interpreted in such an unmediated context is not addressed by Veltman, but the answer will be highly instructive.

As before, I will pursue the answer step by step. Considering the succession of updates note that being ignorant with respect to any background laws means that no nomological updates will occur. Rather, in this scenario, the discourse initial factual update will operate directly on the minimal state. Hence, communicating *John did not drink* ($= \neg p$) will lead to a fully symmetric partition of the available logical space into a factual and a counterfactual domain (again distinguished in the table below by the labels *fact* resp. *ctfct.*). Since in this case there are no laws that could define relations between facts, all (atomic) sentences are fully independent. The bases of the worlds in this context are thus generalized and given by the truth value assignments for p , q , and r , i.e. the worlds themselves. When confronting the counterfactual *If John had drunk, he would have become sick* the retraction state $\mathbf{1}[\neg p] \downarrow \llbracket \neg p \rrbracket$ is thus defined by the values of q and r , essentially leaving all worlds (be that factual or counterfactual) up for consideration. Making

CHAPTER 4. THE ANALYSIS

the counterfactual assumption by updating the retraction state with p again re-establishes the distinction between factual and counterfactual worlds. (The state is thus $\mathbf{1}[\neg p] \downarrow \llbracket \neg p \rrbracket [p]$.) Now, what is different to the original example is that in this nomologically unrestricted case the support test $\mathbf{1}[\neg p] \downarrow \llbracket \neg p \rrbracket [p] \models q$ obviously *fails*:

	p	q	r	$\neg p$	basis	$\downarrow \llbracket \neg p \rrbracket$	p	$\models q$
w_0	0	0	0	fact	$\langle p, 0 \rangle, \langle q, 0 \rangle, \langle r, 0 \rangle$	\checkmark	—	—
w_1	0	0	1	fact	$\langle p, 0 \rangle, \langle q, 0 \rangle, \langle r, 1 \rangle$	\checkmark	—	—
w_2	0	1	0	fact	$\langle p, 0 \rangle, \langle q, 1 \rangle, \langle r, 0 \rangle$	\checkmark	—	—
w_3	0	1	1	fact	$\langle p, 0 \rangle, \langle q, 1 \rangle, \langle r, 1 \rangle$	\checkmark	—	—
w_4	1	0	0	ctfct.	$\langle p, 1 \rangle, \langle q, 0 \rangle, \langle r, 0 \rangle$	\checkmark	1	0
w_5	1	0	1	ctfct.	$\langle p, 1 \rangle, \langle q, 0 \rangle, \langle r, 1 \rangle$	\checkmark	1	0
w_6	1	1	0	ctfct.	$\langle p, 1 \rangle, \langle q, 1 \rangle, \langle r, 0 \rangle$	\checkmark	1	1
w_7	1	1	1	ctfct.	$\langle p, 1 \rangle, \langle q, 1 \rangle, \langle r, 1 \rangle$	\checkmark	1	1

Besides worlds w_6 and w_7 for which both p and q are true, now also worlds w_4 and w_5 (which were in the previous example excluded on grounds of the presupposition of the law $\Box(p \rightarrow q)$) have to be considered. And for these worlds it holds that although they are worlds in which *John drinks* is true, the proposition that *John is sick* is false. (For simplicity's sake the truth values are again restated directly in the respective lines.) Thus, the failure of the support test. This however raises a serious question: how is the addressee B supposed to deal with the speaker A 's utterance of a counterfactual that is so obviously failing its function? How could B even consider A to be rational in this situation? And, turning away from these empirical questions, what does this case mean for Veltman's theory?

The assumption of rationality actually provides with a first solution to this. (Later on, when discussing as a fourth example the discourse *John drank too much. If he had not drunk too much, he would not have become sick* (cf. section 4.2.4, p. 129 ff.) another, competing solution will be formulated.) B can explain A 's behavior to be rational by assuming that A entertains some nomological information so far unbeknownst to B in light of which the support test is rendered successful. I.e.

CHAPTER 4. THE ANALYSIS

in order to make sense of *A*'s utterance, *B* has to assume that *A* knows some law from which derives – when making the respective counterfactual assumption – the counterfactual's consequent as a valid conclusion.

In order to achieve this and to reach a congruent information state between *A* and *B*, *B* has at least two options. The first is direct inquiry in form of *B* challenging the counterfactual and asking *A* to provide the missing nomological information. The other option is less expensive, in that it saves the communicative effort connected to the first strategy. Here *B* simply *accommodates* the presupposed rule from the given counterfactual, taking its indicative correlate as a nomological update.⁹ I.e. *B* abstracts from the counterfactual *If John had drunk any wine, he would have become sick* the rule *If you drink, you will become sick* and updates his or her information state with this nomological proposition. After having done so the support test connected to the counterfactual will no longer fail, but be successful as has been argued for in detail when discussing the first example above. (By accommodating the law $\Box(p \rightarrow q)$ the worlds w_4 and w_5 will be excluded from having to be considered, and by making the counterfactual assumption that *p* it then follows by *Modus Ponens* that *q*: if you presuppose that *drinking will result in sickness* and assume that *John drank*, it derives that *John is sick*.)¹⁰ On this account this strategy is doubly informative: first, the counterfactual turns out to be informative, because *B* gets to learn a new rule by accommodating new nomological information, and second, because by doing so *B* can maintain (or confirm) the assumption that *A* is a rational partner in communication since his or her utterance appears informationally motivated. Thus, in combination both these steps yield a definite profit at minimal costs. (By this also the apparent pragmatic perversity of the discourses is explained away, for, after all, the question is why should anyone

⁹Recall Lewis *rule of accommodation for presupposition*: “If at time *t* something is said that requires presupposition *P* to be acceptable, and if *P* is not presupposed just before *t*, then – *ceteris paribus* and within certain limits – presupposition *P* comes into existence at *t*.” (Lewis (1979b), p. 340) While the notion of accommodation remains rather vague and resembles some kind of *emanatio intellectualis* the issue shall not be pursued any further. Rather I follow Lewis's lead here, who remarks following his statement of the rule of accommodation: “This rule has not yet been very well stated, nor is it the only rule governing the kinematics of presupposition. But let us bear it in mind nevertheless, and move on to other things.” (Lewis (1979b), p. 340)

¹⁰Actually, the table that summarizes the initial example represents also the effect of accommodation discussed here.

CHAPTER 4. THE ANALYSIS

first give some factual information only to immediately follow up with an exposition of what would have been the case if the just communicated fact had not been the case? Since the communicated fact and its counterfact are mutually exclusive either of these bits of information seems irrelevant and not worth the effort of mentioning.)

4.2.3 The third case: the infelicitous # *John drank. He would not have become sick*

Having discussed the case of the felicitous discourse *John did not drink (any wine). He would have become sick*, it suggests itself to contrast this example with its infelicitous counterpart # *John drank (too much wine). He would not have become sick*.

While it was already suggested in the course of the discussion of the *because*-paraphrase of the data (cf. section 1.6, p. 21 ff.) that for the felicity of a discourse the availability of a NinA is not the crucial factor, I will nonetheless discuss in detail what the absence of a NinA (meaning really the (practically pretty much impossible) failure to accommodate a suitable antecedent) would mean for the semantic computation of the counterfactual. The goal is to understand in detail what repercussions would arise from this for Veltman's approach. First I will recapitulate the argument that motivates the introduction of the concept of the negation induced alternative. Then I will discuss the example in question also formally.

An informal recount of the NinA-story

Turning to the concept of the negation induced alternative the starting point was provided by Veltman's observation that certain counterfactuals are not paraphrasable while others are. Recall the initial examples, that are restated here in their simplified form as (3) and (4):

- (3) If John had drunk, he would have become sick.
- (4) If John had not drunk, he would not have become sick.

CHAPTER 4. THE ANALYSIS

While the former can be paraphrased in terms of the discourse given in (5), trying to do so for the latter yields an infelicitous sequence, as attested by (6), although the exact same principles of reformulation underlie this rendering.

(5) John did not drink. He would have become sick.

(6) # John drank. He would not have become sick.

The thesis that in the former case a contextually given element – the negation induced alternative – acts as an implicit antecedent to the counterfactual consequent, while in the latter case no such link to the counterfactual domain is provided by the context (and hence the counterfactual remains incomplete and the discourse thus infelicitous) finds strong empirical support when turning to another observation made by Veltman in connection with the data. The point is twofold and centres around the insertion of the particle *otherwise* in the discourses (cf. Veltman (2005), p. 160). When inserted into (5) the semantics of the discourse does not change.

(7) John did not drink. *Otherwise* he would have become sick.

Since (7) is completely synonymous to (5), the particle *otherwise* is either semantically vacuous, or its semantic contribution to (7) must already be implicit also in its counterpart in (5). Intuitively the case is clear: *otherwise* is certainly not empty. Rather it mirrors the meaning of the preceding initial sentence:

John did not drink. *Otherwise* . . .
= John did not drink. *Otherwise* = *If he had drunk* . . .

Hence, for the simple examples under discussion in which the initial sentence is a negated or non-negated atomic proposition p that is followed by an *otherwise*-sentence, the meaning of *otherwise* can very roughly be given as:

[[*otherwise*]] = If it had not been the case that p . . .

Ascribing this meaning to *otherwise* also explains why its insertion into the infelicitous example (6) actually renders the resulting discourse all of a sudden felicitous (which eventually provides with factual empirical support for the cognitive reality of the NinA) :

CHAPTER 4. THE ANALYSIS

(8) John drank. *Otherwise* he would not have become sick.

If *otherwise* were indeed semantically empty, one would expect the felicity of the example not to change. However, when assuming that *otherwise* acts as an operator that mirrors the semantic content provided by the discourse initial sentence things fall into place: qua its capacity to mirror information given in the preceding context the particle *otherwise* provides with an antecedent to the counterfactual consequent and thus completes the counterfactual conditional. Felicity is thus guaranteed. As the examples in (7) and (8) show, the polarity of the discourse initial context is irrelevant in this. Polarity however is an issue in the cases in which *otherwise* is not inserted into the discourses. Since in the infelicitous example (6) the discourse initial sentence is affirmative, lacking a (sentential) negation, no negation induced alternative is made available that could act as the implicit antecedent to the counterfactual consequent. The context is simply not rich enough to complete the counterfactual. Since it remains defective, the discourse is hence rendered infelicitous. –Without a proper premise the asserted conclusion can not be assessed, and thus does not find ‘support’, to stick with Veltman’s terminology.

The formal repercussions of the absence of a NinA

Point of departure for the formal recount of the infelicitous discourse is a scenario in which the nomological relation between drinking (translated in the following again as p) and sickness (translated as q) is presupposed. Hence, the universe U_S of the state is restricted by the rule *If you drink too much, you will become sick*: $\Box(p \rightarrow q)$, after the minimal state $\mathbf{1}$ has been updated accordingly. This state $\mathbf{1}[\Box(p \rightarrow q)]$ is next updated with the factual proposition *John drank* (as an instantiation of p) expressed by the discourse initial sentence. The state thus reached is $\mathbf{1}[\Box(p \rightarrow q)][p]$ When following up with the counterfactual consequent *He would not have become sick* things become interesting. Trying to embed the consequent (represented by ‘ $\dots \models \neg q$ ’ for its demand of a supporting context) into the context built up so far yields the following situation: $\mathbf{1}[\Box(p \rightarrow q)][p] \dots \models \neg q$.

Prerequisite for the evaluation of the support test put forth by ‘ $\dots \models \neg q$ ’ is the availability of a suitable counterfactual context. Since $\mathbf{1}[\Box(p \rightarrow q)][p]$ designates

CHAPTER 4. THE ANALYSIS

the factual dimension of the discourse it will not do, because the irrealis of the counterfactual consequent requires a counterfactual support.

Also, as has already been discussed at length, the counterfactual complement to the factual dimension of discourse (which is given by $\mathbf{1}[\Box(p \rightarrow q)] \setminus [p]$ (i.e. the set of the (law-abiding) worlds in which *John does not drink*) does not represent the proper set of premises that could fill in the dots in ‘... $\models \neg q$ ’. It is too indiscriminate with respect to the worlds it makes available to fill in the ‘...’ in ‘... $\models \neg q$ ’, containing also worlds that differ from $w_{@}$ more than admissible.

Trying to construct a proper context by applying Veltman’s algorithm however also fails, because it simply can not be executed. Neither the counterfactual update nor even the retraction state $\mathbf{1}[\Box(p \rightarrow q)][p] \downarrow \llbracket [p] \rrbracket$ can be achieved: since there is no salient antecedent available – as if were the case had a NinA been generated by the context – it is for one impossible to determine what proposition is to be retracted from the bases of the factual worlds, and for another it further is not clear what proposition acts as the counterfactual update. This effectively sabotages the whole process of making a counterfactual assumption: the support test put forth by the counterfactual consequent fails, because the ‘...’ in ‘... $\models \neg q$ ’ can not be filled in and the premise of the support test remains undetermined. (Note that strictly speaking, it can be filled in, but with an improper set of premises. The premise set actually consists of the whole universe U_S that does not differentiate between factual and counterfactual worlds. Which means that in the absence of a NinA the required domain of the support test is not available.)

4.2.4 The fourth case: *John drank. If he had not drunk, he would not have become sick*

Having discussed the infelicitous # *John drank too much. He would not have become sick*, I will next concentrate on its non-elliptic, and hence felicitous counterpart *John drank. If he had not drunk, he would not have become sick*. It will turn out that also this discourse is highly interesting. To understand the significance of the example the discussion up to this point shall first be briefly summarized.

CHAPTER 4. THE ANALYSIS

The first example to be semantically analyzed was the original discourse *John did not drink any wine. He would have become sick* containing an implicit antecedent and the presupposition of the law that *Drinking too much alcohol will cause sickness*. The second point of discussion was provided by the same discourse but this time its interpretation was radicalized insofar it was not aided by the assumption of the nomological presupposition. We saw that whereas in the first, presuppositional case the counterfactual was trivially true, in the second case (that lacks the presupposition) the counterfactual was trivially false. In explaining how interpreters nonetheless could – and, more often than not, do – accept the counterfactual conditional as true, a general assumption of rationality in communication was pivotal. While in both these cases the discourses are elliptic, a consideration of their non-elliptic analogue in which the antecedent appears explicitly is not necessary, since the implicit antecedent of a discourse is actually identical to the antecedent in an explicit discourse. Both variants contain the same semantic object in the *if*-clause of the counterfactual conditional.

In the third variant # *John drank too much wine. He would not have become sick* that was next attended, the discourse initial sentence of the discourse does not contain a negation, but is affirmative, and is thus lacking the potential to induce a NinA as an antecedent to the counterfactual consequent. Hypothetically it was then discussed what the absence of an antecedent of the counterfactual would mean for the evaluation of the consequent.

Next, I will finally concentrate on the variant of the third example in which the antecedent is explicitly given and the discourse thus felicitous: *John drank (too much wine). If he had not drunk (too much wine), he would not have become sick*. For this fourth case I will additionally assume that the discourse is embedded in a presuppositional scenario: it has been established as a law that *Drinking will cause sickness*. With this the example seems to be built by and large analogously to the first example, and thus may appear to be equally trivial. But we will see that this is not the case. Ultimately the presupposition and the information given by the discourse are mutually logically independent. Thus, the example does not run

CHAPTER 4. THE ANALYSIS

parallel to the first example, but rather runs parallel to the second example that was concerned with the radical interpretation of counterfactuals. In both these cases the counterfactual is to be interpreted without a relevant nomological backbone that could deliver the axiomatic foundation for an assessment of the soundness of the counterfactual. This means that also for the example to be discussed now the question arises why interpreters do not take this kind of counterfactual to be *a priori* false.¹¹ As for the other example Veltman does not deliver an answer also for this case. Some contemplation will however reveal that there is more than one solution to this problem. While one solution has already been formulated at the end of the first chapter (cf. p. 28 ff.) some further consideration will help to formulate at least one more answer to this question. This (competing) answer will again be based on pragmatic strategies. Taken together both these proposals eventually confirm a point about counterfactuals that Veltman puts as follows:

It has often¹² been noted that general laws play a special role in the interpretation of counterfactuals. [...] In making a counterfactual assumption we are not prepared to give up propositions we consider to be general laws. (Veltman (2005), p. 166)

We will see that in case the context of interpretation is nomologically vacuous, rational agents may devise strategies for an *ad hoc*, but principled construction of the necessary information. Keeping this in mind, I will turn now to the discussion of the fourth example.

As before I will initially assume for the scenario that the law *If you drink too much, you will become sick* can be presupposed (although we will see that this presupposition is ultimately irrelevant, as has already been mentioned). This in turn means that the bases of the worlds under consideration are the same as in the first example. It is only the factual updates in which this fourth example is

¹¹Assuming that the proposition *If you do not drink too much alcohol, you will not become sick* has the status of a law would of course alter the picture. But this would stretch the understanding of what characterizes a ‘law’ rather far, since there are more causes to becoming sick imaginable than just drinking.

¹²Veltman refers at this place in a footnote to Pollock by remarking: “A prominent example is John Pollock, who has stressed the point in all his writings on counterfactuals since Pollock (1976).” (Veltman (2005), p. 166, footnote 13)

CHAPTER 4. THE ANALYSIS

different from the first one. Indeed, the situation is reversed: what had previously been designated as fact now turns to counterfact and the other way around.

Point of departure now is the update with the proposition p , i.e. the discourse initial non-negative sentence *John drank* being communicated. The respective column in the table below represents this first factual update. Next, the explicitly given *if*-clause of the counterfactual conditional is uttered. This triggers the construction of the retraction state and the counterfactual update. The former is given in the column headed by $\downarrow \llbracket p \rrbracket$. As before, all nomologically sound worlds are available in this state, while those worlds that do not comply with the rule remain exempt from any consideration. The counterfactual assumption that $\neg p$ (standing for *John did not drink*) introduced by the corresponding update with the *if*-clause of the counterfactual is established in the following column. Again this differentiates the counterfactual worlds from the factual worlds. The latter (w_6 and w_7) may be ignored and are thus marked ‘—’ at this stage in the table. Up for consideration remain only the counterfactual worlds w_0 , w_1 , w_2 , and w_3 . In these worlds the proposition *John did not drink* is true. The conception of counterfactual conditionals as support tests requires that in these worlds the counterfactual consequent *He_{John} would not have become sick* must be true, if the counterfactual *If John had not drunk, he would not have become sick* as a whole is to be true. For most people the first intuition regarding the truth or falsity of the counterfactual in the given scenario is probably (at least when assuming the truth of the discourse initial sentence) that it is indeed true. However, when consulting the table to validate the success of the support test, it turns out that it fails, since, as specified again directly in the table, while $\neg p$ is true in all worlds under consideration, $\neg q$ is not true in all these worlds. In fact, in worlds w_2 and w_3 it is false:

CHAPTER 4. THE ANALYSIS

	p	q	r	$\Box(p \rightarrow q)$	p	basis	$\Downarrow \llbracket p \rrbracket$	$\neg p$	$\models \neg q$
w_0	0	0	0	\checkmark	ctfct.	$\langle q, 0 \rangle, \langle r, 0 \rangle$	\checkmark	1	1
w_1	0	0	1	\checkmark	ctfct.	$\langle q, 0 \rangle, \langle r, 1 \rangle$	\checkmark	1	1
w_2	0	1	0	\checkmark	ctfct.	$\langle p, 0 \rangle, \langle q, 1 \rangle, \langle r, 0 \rangle$	\checkmark	1	0
w_3	0	1	1	\checkmark	ctfct.	$\langle p, 0 \rangle, \langle q, 1 \rangle, \langle r, 1 \rangle$	\checkmark	1	0
w_4	1	0	0	—	—	—	—	—	—
w_5	1	0	1	—	—	—	—	—	—
w_6	1	1	0	\checkmark	fact	$\langle p, 1 \rangle \langle r, 0 \rangle$	\checkmark	—	—
w_7	1	1	1	\checkmark	fact	$\langle p, 1 \rangle \langle r, 1 \rangle$	\checkmark	—	—

While this outcome is (as before with the case discussed in section 4.2.2 above) of course rather puzzling, I will however not take it as evidence against the correctness of Veltman’s theory. Rather I will take it to shed light on certain peculiarities connected to the semantics of counterfactual conditionals and certain cornerstones of the reasoning employed in the example that go unmentioned by Veltman. Actually, it might be argued that in Veltman’s system basic intuitions are indeed captured: although interpreters at first might naively tend to agree with the counterfactual, a second consideration may lead to dispute the consequent and may raise concern about the validity of the counterfactual. How this might come about, can be explained when considering the crucial worlds w_2 and w_3 in detail. Recall that in these worlds it is true that *John did not drink*, but false that *He did not become sick*, i.e. these are worlds in which John did become sick. Imagining counterfactual circumstances that facilitate such a scenario is easy. Just think for instance of a situation such as a party that John attended, and at which he did not drink, but got sick due to food poisoning, because the chicken sandwiches he ate were spoiled.¹³ In this case the counterfactual conditional *If he had not drunk, he would not have become sick* is indeed false. This clearly contradicts the initial naive intuition of the counterfactual being true (i.e. the case in which the worlds w_2 and w_3 are for some reason ignored in the assessment of the counterfactual). But then the question is, how could this understanding come about to begin with? How could it be explained that the first interpretation proceeds selectively with

¹³Considerations like this finally justify the presence of the third proposition r in the table: r captures possibilities like the one mentioned. The point actually directly relates to the problem of counterfactual conditionals, as formulated for instance in Goodman (1947).

CHAPTER 4. THE ANALYSIS

respect to the logical space? –After all, the counterfactual is at first arguably not disputed by most interpreters, although it fails in terms of the support test.

Rationality delivered the key to the first answer to this question as formulated in section 4.2.2. Briefly summarized the argument was that when presuming that the speaker is rational in his or her utterance of the counterfactual *If John had not drunk any wine, he would not have become sick* the speaker has good reason to accommodate the corresponding nomological proposition $\Box(\neg p \rightarrow \neg q)$. This leads to the exclusion of worlds w_2 and w_3 that initially led to the failure of the support test. –Which in turn renders the support test successful and the counterfactual true.¹⁴ But there is more to be said about this point. Hence, I will discuss this issue further. Also in the discussion to follow pragmatic reasoning will motivate the case. But this time in a more elaborate manner.

More on radical interpretation of counterfactuals: conditional strengthening

Central to the line of reasoning that summarizes an argument given by von Fintel (2001) is a quantity implicature¹⁵ that plays upon the locutionary relations and informativeness of the possible antecedents of the counterfactual.

Starting point is the following general assumption about the pragmatics of uttering an antecedent in conditional sentences.¹⁶ Whenever an antecedent q is more informative than an alternative antecedent p (meaning that q entails p), then one

¹⁴This requires of course a rather strong commitment on the interpreter's side with respect to the speaker's reliability. A more skeptical mind would probably not follow this strategy.

¹⁵The following quote from von Fintel (2001), p. 11 explains the notion 'quantity implicature' as follows: "It is usually assumed that quantity implicatures are based on so-called Horn-scales, scales made up of progressively stronger statements. If a statement below the top of the scale is asserted, there will (if everything else works right) be a quantity implicature that the speaker is not in a position to assert any of the stronger elements on the scale." Since the implicature under discussion here is concerned with the first Gricean Maxim of Quantity that reads *Make your contribution as informative as is required (for the current purposes of the exchange)* (as opposed to the second Maxim of Quantity that states *Do not make your contribution more informative than is required* (cf. Grice (1975), p. 45), it might also be called (as Matsumoto for instance does, cf. Matsumoto (1995), p. 21) a *Quantity-1 implicature*.

¹⁶I adopt here a formulation given in von Fintel (2001), p. 1 regarding the pragmatics of assertion. Since a conditional's antecedent is *assumed* rather than *asserted* with its utterance the formulation is adjusted accordingly, using the notion 'utter/assume' instead of 'assert'.

CHAPTER 4. THE ANALYSIS

may assume that:

- (1) *If q were the case, p would not have been uttered/assumed.*¹⁷

Applying this to the present example is straightforward: while p designates the actually uttered/assumed antecedent *John did not drink*¹⁸, the propositional variable q stands for the conjunction of all possible ‘reasons’ that derive the consequence *John would not have become sick*. I.e. the (exhaustive) set of possible answers to the question *When would John not have become sick?* Namely, *(If) John had not drunk* AND *(if) he had not eaten the spoiled chicken sandwich* AND *(if) he had not caught the Noro-virus* AND *(if) he had lived healthily* AND *(if) ...* etc. (For this reason I will from now on refer to the set of possible answers (i.e. possible antecedents) by p^{AND} , instead of q that was used so far.¹⁹) Since however, p has been uttered/assumed, it follows (as von Fintel points out²⁰) from (1) and the assumption of the truth of the uttered antecedent p (the proposition that *John did not drink*) by *Modus Tollens* that p^{AND} is not true. I.e. none of the alternatives p_n to p occurring in p^{AND} obtains.

¹⁷While von Fintel’s take here differs in its formulation from the UTTER/ASSUME-operator that was adapted from Krifka and employed in the discussion earlier (cf. section 1.7.1, page 29), the outcome of both approaches is ultimately very similar in that in the end the conditional is strengthened to a bi-conditional.

¹⁸A proper translation of the uttered/assumed antecedent would have to observe the expressed negation. For simplicity’s sake I take p (and the p_n to be introduced in a moment) however to designate *any* premise, disregarding its specific polarity. The truth values I am referring to in the argument are thus to be understood independent from the specific form of a premise.

¹⁹In my emphasis of p^{AND} I take conjunction to have at this point of the discussion intuitive precedence over the equivalent formalization using negation and disjunction. The matter is however not essential, since a transformation into an equivalent (disjunctive, resp. conjunctive) normal form is always possible. In either case the idea is that it holds that none of the possible alternative antecedents must obtain.

²⁰cf. von Fintel[2001], p. 12 - 13, where he writes: “The hearer is not reasoning about a long (or infinite) list of possible antecedents. The hearer is simply reasoning that if there were an antecedent r (other than p) such that *if r , q were true* [besides *if p , q* , D.B.], the speaker would have added this conditional to the assertion. But we can appeal to quantity at this point: the reason why the speaker would have added such a conditional to the one actually asserted is that the conjunction would have been a statement that should have been asserted (because it gives more of the required information).” (Underlining in the original, D.B.) Since von Fintel discusses alternative conditionals *if r , q* and *if p , q* , where I discuss alternative antecedents to the consequent he may employ the notion of *assertion*, where I have to use *utter/assume*.

CHAPTER 4. THE ANALYSIS

Seeing that p^{AND} entails p and is hence more informative (and builds a Horn-scale) is easy: since p acts as a conjunct in p^{AND} , it holds by the definition of conjunction that whenever p^{AND} is true, also p must be true. The question how one can guarantee that indeed all other premises p_n besides p are false appears to be less trivial, since p^{AND} is possibly infinite. The answer that is developed in von Fintel (2001) (with reference to Cornulier (1983) and Matsumoto (1995)) refers to quantity as the crucial element and plays upon exhaustivity in an peculiar manner: given that for any single conjunct p_n occurring in p^{AND} such that $p_n \neq p$ the conjunction $p \wedge p_n$ is more informative than p alone, by asserting p it follows from (1) that p_n must be false. Hence one can conclude that for no premise p_n it holds that it presents a true alternative to p –which essentially exhausts the set of possible antecedents.²¹

As a result of the reasoning so far this means that – since all other possible alternatives p_n are to be excluded – p acts as the *only* valid premise to the counterfactual consequent. But then it follows that if p were not the case, the consequent would not derive. Both points brought together however yield conditional strengthening, i.e. the interpretation of the conditional as a bi-conditional.²² From

²¹ At this point one has to be cautious. The following problem arises that goes unnoticed in von Fintel (2001) (and the literature referred to therein): since it is concluded that any p_n is false (as for instance the (counterfactual) premise that *John did not catch the Noro-virus*) it seems to follow immediately that *John caught the Noro-virus* is true. This however is an unwelcome over-determination of the actual situation. The only ‘reason’ for John’s being sick (presupposing the truth of the counterfactual consequent’s complement) mentioned is that *John drank* –and nothing else. Hence, all other possible ‘causes’ for John’s sickness are to be excluded. And although this kind of selective understanding arguably happens, the question how this can be accounted for is not quite clear to me, though. While one might think of some form of the *closed world assumption* as one possible key to an explanation, the main problem is in any case that not just all other hypothetical premises, but also their negations have to be false. –Which is impossible in a bivalent account. (In Reiter (1978) the closed world assumption is informally paraphrased as follows: “[...] it is natural to explicitly represent only positive knowledge and to assume the truth of negative facts by default. For example, in an airline data base, all flights and the cities which they connect will be explicitly represented. Failure to find an entry indicating that Air Canada flight 103 connects Vancouver with Toulouse permits one to conclude that it does not.” (Reiter (1978), p. 56) A later informal elaboration in Reiter (1988) states: “The idea is that the knowledge base is viewed as completely representing all the positive information about some world. Any ground atomic fact not so represented is taken to be false.” (Reiter (1988), p. 102)

²²As von Fintel points out this observation is due to Geis and Zwicky (1971), who use the term *conditional perfection*, instead of the notion of *conditional strengthening* that is used here following von Fintel. The objection that besides p^{AND} and p also *iff* and *if* build a Horn-scale,

CHAPTER 4. THE ANALYSIS

$p \rightarrow \text{counterfactual consequent}$

and

$\neg p \rightarrow \neg \text{counterfactual consequent}$

it derives that

$p \leftrightarrow \text{counterfactual consequent}$

since the following equivalence is a propositional tautology:

$$((P \rightarrow Q) \wedge (\neg P \rightarrow \neg Q)) \Leftrightarrow (P \leftrightarrow Q)$$

On ground of this line of reasoning worlds w_2 and w_3 are eventually excluded from the states' logical space, since in these worlds the truth value assignments of antecedent and consequent differ. Conditional strengthening thus formally guarantees that the support test succeeds, because in the remaining worlds w_1 and w_2 both antecedent and consequent turn out to be true. But it has also welcome corresponding conceptual repercussions. Recall w_2 and w_3 are worlds in which John does not drink, but still becomes sick. Thus excluding these worlds by conditional strengthening actually means to exclude all other possible explanations of John's physical condition besides the one asserted in the antecedent.

since *iff* entails *if* – which would sabotage the whole argument, since an implicature based on this observation would derive anything but conditional perfection – is countered by von Fintel with reference to Matsumoto (1995) (cf. von Fintel (2001), p. 14). Matsumoto argues that the decisive criterion for a Horn-scale is that the items building the scale share the same monotonicity properties, and that this is not the case for *iff* and *if*: “*If and only if* clauses display a phenomenon parallel to non-monotone NPs: this clause does not allow its consequent clause to be strengthened or weakened *salva veritate* [...]. This contrasts with *if* clauses, which exhibit a phenomenon parallel to monotone increasing NPs [...]. An *if* clause allows its consequent to be weakened but not strengthened *salva veritate* [...]. These tests show that *if and only if* clauses are not scalar in nature, while *if* clauses are scalar (on a positive scale). Therefore these two clauses do not form a Horn scale.” (Matsumoto (1995), p. 47 - 48) In contrast to this p^{AND} and p share the same monotonicity properties, since the connective *and* does not alter the monotonicity properties of its (atomic) conjuncts.

4.2.5 Generalizing the point

Note that the discussion is of general importance and describes a major point about the semantics of counterfactuals. In its result it clarifies certain rescue mechanisms regarding the (radical) interpretation of counterfactuals in case they do not have a nomological background against which they could be assessed.

The described strategy of strengthening applies also to discourses like *John did not have surgery. If he had had surgery, he would not have died* which was brought to attention at the end of the first chapter. Also for this example is the nomological background empty: world knowledge tells us that the interpretation of the counterfactual in *John did not have surgery. If he had had surgery, he would not have died* can not take recourse to background laws. Having surgery simply does not guarantee the well being of the patient. This means for the counterfactual that it does not describe a traceable inference, but rather simply stipulates a certain relation between propositions. Hence, in principle also this counterfactual is trivially *false*. That interpreters do nonetheless not universally reject the conditional may be motivated by lines of pragmatic reasoning as described in 4.2.2 and 4.2.4.

4.3 The solution to Veltman's puzzle: felicity conditions for elliptic counterfactuals

What remains finally to be answered are the questions connected to the origin and role that exhaustivity plays in the felicity of the examples. As with *John drank. If he had not drunk, he would not have become sick* also in case of *John did not have surgery. If he had had surgery, he would not have died* exhaustivity was assigned a decisive role in the discussion. In case of the former – just above – with respect to its semantic acceptability (when conceiving the meaning of a conditional as a support test, cf. section 4.2.4, p. 134 ff.), and in case of the latter already in the first chapter (cf. section 1.7, p. 25 - 28) in terms of its pragmatic felicity.

In the first chapter it was initially assumed that it is negation in the consequent that is responsible for the introduction of an exhaustivity constraint on the set of

CHAPTER 4. THE ANALYSIS

possible premises to the counterfactual consequent. And while exhaustivity indeed figures as a decisive factor in the account of the data, the felicitous examples introduced in footnote 26 on page 26 that are repeated here in (9), (10), and (11) cast a very strong doubt about the role and responsibility of negation in this matter.

- (9) John did not drink any alcohol. He would not have passed the impending drug test.
- (10) The pilot took the compass with him. He would not have found his way home.
- (11) John won the race. He would not have celebrated.

All these examples do contain a negation in the consequent, but none of them are infelicitous. Hence, a better lead to the solution has to be followed.²³

The solution is to consider what renders felicitous examples like (9), (10), and (11) different from those that are infelicitous. So, reconsider an example already discussed in the first chapter:

- (12) # John did not have surgery. He would not have died.

The key to the problem is actually provided by the procedure that allowed to identify the exhaustivity constraint to begin with: the question which conditions would have to obtain in order to render the consequent true. This is the underlying question under discussion (that also is pivotal in von Fintel's account of conditional strengthening²⁴).

In case of # *John did not have surgery. He would not have died* the possible answers built a conjunctive array of propositions. In case of the examples (9) -

²³Also the fact that the substitution of a negative consequent by a non-negative paraphrase does not always yield a change in acceptability causes doubt about the role of negation. The following example illustrates the point. Not just the discourse # *John did not have surgery. He would not have died* is infelicitous, but also a near synonymous non-negative counterpart is infelicitous: # *John did not have surgery. His condition would have remained unchanged.*

²⁴Cf. von Fintel (2001), p. 17, where he remarks: "The idea that we can now explore is that perfection will be triggered if the conditional is asserted as an answer to a question eliciting an exhaustive list of sufficient conditions for the consequent."

CHAPTER 4. THE ANALYSIS

(11) the respective answers however turn out to be combined disjunctively:²⁵

(9)': When does one not pass an impending drug test?

In case one drinks alcohol before the test, or smokes pot, or takes cocaine, or ...

(10)': When do you not find your way home?

If you do not have a compass, or if you have the wrong map, or if it is dark, or ...

(11)': When does one not celebrate?

If X happens, or Y happens, or Z takes place, or ...

Whenever the antecedent of a counterfactual remains implicit, a conjunctive premise set – as opposed to a disjunctive premise set – can not be exhausted, hence the discourse is rendered infelicitous. The semantic demands for the evaluation of the consequent (namely to exhaust the set of *all* possible premises) clash in case of a conjunctive premise set with the meaning of the subjunctive, resp. the concept of assumption and the pragmatics of assuming which demand that *anything* is possible.

The impact and importance of the distinction between conjunctive versus disjunctive premises is highlighted also in the explanation of another phenomenon that otherwise would figure as a permanent thorn in the side of the initially formulated NinA-proposal: why is it that also discourses without a NinA in the context may be acceptable, while others which are equipped with a NinA are not? Examples were:

²⁵Translating these properties into more classical terms, the latter set of answers characterizes a set of *sufficient* conditions, while the former designates *necessary* conditions for the truth of the consequent. The logic that underlies the relation between disjunctive and conjunctive premise sets and the occurrence of negation is given by the following equivalence:

$$(p_1 \vee p_2 \vee p_3 \vee \dots p_n) \leftrightarrow q \quad \Leftrightarrow \quad (\neg p_1 \wedge \neg p_2 \wedge \neg p_3 \wedge \dots p_n) \leftrightarrow \neg q$$

The intuition that corresponds to this tautology is twofold: First, not only follows from the truth of at least one of the (sufficient) premises the truth of the consequent, but also from the truth of the consequent follows the truth of at least one of the premises. And second, whenever one of the premises that is sufficient for the truth of the consequent is indeed met, then it can not be the case that also one of the premises sufficient for the negation of the consequent obtains. (And vice versa.)

CHAPTER 4. THE ANALYSIS

(13) John had surgery. He would have died.

vs.

(14) # John did not have surgery. He would not have died.

Turning to the QUD delivers the answer to this question. The logical structure of the premise set of (14) is that of a conjunction –which, as has been argued at length, can not be exhausted by an implicit antecedent (even though a NinA is available). On the other hand, the answer to the QUD of (13) determines the premise set of the counterfactual to be disjunctive. Hence, it is already sufficient that a single premise is instantiated for the inference described by the (then completed) counterfactual to go through. And although the context is not enriched by a NinA, the discourse initial proposition is arguably salient enough to facilitate the task: its negation (which is procedurally made available due to the dichotomous nature of the update) may act as an implicit antecedent.

While this does not contradict the empirical reality of the concept of the negation-induced alternative, its importance for the account of the data's felicity is of course relativized. But considered the gain in theoretical integrity and completeness for the analysis this can be gotten over.

4.4 On the semantic equivalence of counterfactual discourse and its *because*-paraphrase

What finally shall be addressed is a point that was already highlighted in chapter 1 (and that proved instrumental in the identification of the exhaustivity constraint). It was the observation that the discourses that give rise to Veltman's puzzle have a *because*-sentence paraphrase. The availability of this paraphrase will now lie at the center of attention.

Although the concern is (in a narrow sense) ultimately trivial, since whenever one expression is synonymous to the other the involved meanings have to be identical, the discussion will nonetheless prove to be fruitful. It not only completes the picture and allows to generalize the account, but will give raise to a refinement of the semantics of the connective *because* that was sketched in Lewis (1973a).

The point of departure for the following considerations is delivered by Veltman's original discourse as introduced in the first chapter on page 5 as example (2) (repeated here as (15)) and the corresponding *because*-paraphrase which is repeated here as (16).

(15) John did not drink any wine. He would have become sick

(16) Because he would have become sick, John did not drink any wine

Although the synonymy of the data is entirely intuitive – and thus the semantic relation between (15) and (16) absolutely elementary – it is not so clear how this actually comes about. This will be clarified first.

4.4.1 First observations regarding the synonymy of the data

The overtly given semantic objects

When reconsidering the data a first observation is that the sentence-like²⁶ constituents in both examples are identical. The sentence *John did not drink* not only acts as the initial sentence of the discourse, but also figures as the main clause of the *because*-paraphrase. This holds also for the other discourse constituent resp. sentential element: the counterfactual consequent *He would have become sick* appears as the subordinate clause of the *because*-sentence as well as the second sentence of the discourse. A slight change in linear order highlights this coincidence:

- (17) John did not drink any wine. He would have become sick.
- (18) John did not drink any wine, because he would have become sick.

The implicit semantic elements

Considering the obvious identity of the involved overt constituents, an explanation of the near synonymy of the data is immanent. Since the meaning of the proposition *John did not drink* is in either example completely unproblematic, only the semantics of the subordinate clause of the *because*-sentence seems to require some more attention. After all, it can not be interpreted out of context, but its meaning can only be determined once a proper context is provided. It stands to reason to assume that – just as is the case for the counterfactual discourse – also in the case of the *because*-sentence the required context is provided by an implicit antecedent, and that also the *because*-sentence is in fact elliptic. Indeed, this intuition finds support in the following paraphrases of the *because*-clause. In these the otherwise implicit element is made explicit:

- (19) Because he would have become sick if he had drunk any wine, John did not drink any wine.

²⁶Since the subordinate clause (= the counterfactual consequent) is a propositional element, I will refer to it as a ‘sentence-like’ or ‘sentential’ constituent.

CHAPTER 4. THE ANALYSIS

- (20) John did not drink any wine, because if he had drunk any wine he would have become sick.

What these variants further suggest is to assume that the implicit antecedent (*If*) *John had drunk any wine* has its origin in the main clause *John did not drink any wine*, or, more specifically, in its potential to induce a salient implicit contextual element. –For which the NinA-account would in turn readily provide an explanation.

More evidence for the context-dependency of the subordinate clause in *because*-sentences: discourse behavior and acceptability

That the NinA-account may indeed be of relevance to the *because*-data is implied by a number of parallels in discourse behavior and acceptability between the *because*-paraphrase and the original counterfactual discourse. Recall that examples like (21) below gave rise to the initial naive NinA-account that predicted the unacceptability of a discourse in case the initial sentence lacks a sentential negation (thus failing to provide with a negation induced alternative):

- (21) # John drank. He would not have become sick.

Constructing the analogue *because*-sentence yields the same result regarding the acceptability of the clause:

- (22) # John drank, because he would not have become sick.²⁷

As has already been observed in connection with the elliptic discourses, the insertion of a discourse mirror like *otherwise* will however render an infelicitous *because*-sentence acceptable:

- (23) John drank, because otherwise he would not have become sick.

²⁷Note that there is a secondary interpretation under which the example is felicitous. It is the reading that *John drank in order to become sick* that is available also for the counterfactual discourse.

CHAPTER 4. THE ANALYSIS

Indeed, as with the counterfactual discourse, the insertion of *otherwise* is feasible also in the complementary *because*-sentence that contains a negation in the main clause but no negation in the subordinate clause:

(24) John did not drink, because otherwise he would have become sick.

Sentences like this provide again strong evidence for the empirical reality of a contextually given implicit negation induced alternative being factually present in the data. The argument is just as with the discourses. It shall nonetheless be rephrased here. Contrast example (24) above with example (25) below:

(25) John did not drink, because he would have become sick.

Intuitively both examples are complete synonyms. This however means that either the expression *otherwise* is semantically vacuous, or, alternatively, that the semantic content distributed by the discourse mirror *otherwise* is implicitly given as a hidden antecedent in example (25). Not only firm semantic intuitions, but also theoretical considerations suggest to conclude to the latter: if one were to assume that *otherwise* is meaningless one would run into considerable trouble explaining the remedying effect that the insertion of the expression into examples like (22) has with regards to its felicity. When assuming instead that *otherwise* figures as a semantic connector that operates on the preceding context these difficulties are however avoided. Moreover, this conception would indeed give a straightforward account of the phenomenon.

Summary

Summarizing these considerations it turns out that counterfactual discourse and *because*-sentence differ in only one (overt, lexical) element. This is the connector *because* itself. All other material is identical. –Be that the sentential constituents on the surface, or the implicitly given semantic unit in form of the negation induced alternative that acts as the hidden antecedent to the *would have* + Past Participle clause. Hence, the pivotal element in order to complete the picture is the meaning of the expression *because* itself. It is the key to understanding the source of the synonymy of the data. Focusing next on the semantics of *because* provides

CHAPTER 4. THE ANALYSIS

not only further, independent, support of the idea of an implicit, contextually induced element, but moreover helps to finally account for the special semantic relationship between the elliptic counterfactual discourse and its *because*-paraphrase.

4.4.2 The semantics of *because*

Turning to the semantics of *because* the following truth condition captures the most basic intuitions.²⁸ It comprises a factual and a counterfactual element: An (indicative) sentence *p, because of q* is true, whenever both the propositional arguments *p* and *q* are true and it is further presupposed that *if q had not been the case, then p would not have obtained*. Or, somewhat more formally²⁹

$$\llbracket \text{because} \rrbracket (q_{\text{subordinate clause}}) (p_{\text{main clause}}) = 1$$

iff

$$(i) \quad p = 1 \text{ and } q = 1$$

and it is presupposed that

$$(ii) \quad \neg q \Box \rightarrow \neg p$$

Subsequently I will elaborate on each component of the given truth condition separately. Doing so, it will turn out that part (i) is instrumental when it comes to further motivating the claim that the main clause provides with an implicit contextual filler for the subordinate clause and that the *because*-sentences of concern here are on the surface elliptical. Part (ii) on the other side will be essential for the final clarification of the semantic relation between counterfactual discourse and *because*-paraphrase. In preparation of these points first however some basic semantic intuitions that underlie (i) shall be retraced.

²⁸The formulation is similar to that given in Hsieh (2009), who attributes his rendering to Lewis (1973a) and Dowty (1979) (who in turn refer to (among others) Hume (1999) (Lewis) and von Wright (1963) and von Wright (1968) (Dowty)). Following Hsieh (2009) I will occasionally refer to the given truth condition as a ‘Lewis-style semantics of *because*’ (cf. Hsieh (2009), p. 66).

²⁹To facilitate orientation I annotate the propositional constituents with subscripts designating their syntactic status.

CHAPTER 4. THE ANALYSIS

A brief reflection on the meaning of *because*: the assertive elements as given in part (i) of the truth condition

In how far the meanings of its sentential constituents contribute to the truth conditional semantics of a *because*-sentence as a whole as specified in part (i) of the condition above shall be highlighted next. As already indicated, these considerations will be rather basal. Yet, they help to get acquainted with the data and they also help to sharpen the most elementary intuitions. The first point of attention will lie with the semantic contribution of the main clause. In order to see the import that the main clause has on the meaning of the complex clause consider a scenario in which the main clause is false. It turns out that the sentence *John did not drink because he would have become sick* is false already when the clause *John did not drink* happens to be false. Next consider a scenario in which the subordinate clause is false. Things behave similarly here: the *because*-sentence as a whole is false if it turns out that the ‘reason’ provided by the subordinate clause does not match the facts. E.g. in case the fact that John did not drink is due to him having to drive, rather than him fearing to become sick. Also in this case the sentence *John did not drink, because he would have become sick* is rendered false.

Although these meditations are intuitively compelling and provide firm confirmation for the first part of the semantics of *because*, one question however arises. The question connects to the last point regarding the semantic contribution of the subordinate clause. Speaking of truth conditions, the problem is how one could ascribe at all a truth value to the subordinate clause? After all, as has already been emphasized a number of times, the sequence *He/John would have become sick* is not interpretable out of context: it lacks a truth value in isolation. Only in the context of an appropriate (i.e. irreal) antecedent can the meaning of the counterfactual consequent be evaluated.

Sketching an alternative formalization for the semantics of *because*

In the account given by Veltman this can be rendered more precisely as follows: the *would have* + Past Participle-complex *He would have become sick* may be considered to be true, whenever there is a contextually given set of premises ac-

CHAPTER 4. THE ANALYSIS

cessible that supports the consequence that *He would have become sick*. Put more formally (along Veltmanian lines) the consequent is true whenever it is embedded in a cognitive state S such that it holds that the proposition *He would have become sick* is *accepted* in S , as Veltman would put it:

$$S \models \llbracket \textit{He would have become sick} \rrbracket$$

Although the exact nature of the informational context of the cognitive state is totally undetermined, at least one fact is distinctly given. This is the content of the *because*-sentence's main clause *John did not drink*³⁰ that acts as an individual, factual update to the state S . Since for the *because*-sentence to be true neither sentential constituent must be false, it suggests itself to assume hence that the following must hold:

$$S[\textit{John did not drink}] \models \llbracket \textit{He would have become sick} \rrbracket$$

Yet, the specific constellation of sentence moods requires to modify the formalization of the supposed *support*-relation. The indicative main clause (that qua being in the indicative acts as a factual update) can only indirectly 'support' the *would have* + Past Participle subordinate clause. Indeed, the counterfactual consequent has to find a counterfactual support. In case of the update with the proposition *John did not drink* this is provided by the negation induced alternative to the given update. The NinA is indeed the most salient counterfactual proposition available. Hence, for the cognitive state in question it ought to hold:

$$S[\textit{NinA(John did not drink)}] \models \llbracket \textit{He would have become sick} \rrbracket^{31}$$

Put like this, the first outline of an alternative formalization of the meaning of the expression *because* gains contour. But in addition also the thesis that it is the main

³⁰For simplicity and brevity I continue to leave the complement phrase *any wine* implicit.

³¹This is a simplification. Under the provision that maximal similarity to the actual world is to be maintained the full machinery of Veltman's system is to be applied. The eventual update with the NinA of the main clause *John did not drink* – which completes the counterfactual conditional implicit to the *because*-sentence – involves generating the retraction state. Hence the actual state that ought to support the proposition *He would have become sick* is $S[\textit{John did not drink}] \downarrow [\textit{John did not drink}] [\textit{John drank}]$.

CHAPTER 4. THE ANALYSIS

clause of the *because*-sentence that provides the contextual filler for the counterfactual consequent is fleshed out a bit more. The assumption of an extreme scenario clarifies the point. For this assume that the cognitive agent is linguistically fully competent, but in a state of total ignorance, i.e. the minimal state. When this agent faces the sentence *John did not drink, because he would have become sick*, the update with the main clause is the only resource that allows to differentiate between the factual and the counterfactual dimension of discourse. The main clause thus provides with the only contextually salient counterfactual alternative that could act as an antecedent to the counterfactual consequent. Indeed, it is this context that guarantees the interpretability of the subordinate clause.

Summary

Summarizing the discussion so far it provided with a number of informal arguments regarding the synonymy of an elliptic counterfactual discourse and its *because*-paraphrase. It was pointed out that both kind of data are elliptic on the surface and involve identical semantic objects. Trivially, this contributes to the semantic equivalence. However, important questions still remain open if one is to fully account for the availability of the *because*-paraphrase. The most important question is to address the semantic contribution of the presupposition given in the second half of the Lewis-style postulate of the meaning of *because* in relation to the semantics of the counterfactual discourse. After all, it is rather intriguing that this – as it turned out actually genuine – part of the meaning of *because* seems to be semantically rather non-essential in the relation between the discourse and its rendering as a *because*-sentence. –At least when considering the other semantic objects involved in the data. How could it be that this apparent semantic surplus of the *because*-sentence does not contrast with the semantics of the discourse? Does the discourse perhaps contain the same presupposition? But if so, what is its trigger? Or is the presupposition otherwise entailed by the discourse? When next these questions are answered the source and the possibility of the peculiar semantic relation between the different kind of data will finally be clarified.

4.4.3 On the presupposition of *because* and the semantic relation between elliptic counterfactual discourse and the *because*-paraphrase.

The point of departure for the discussion to follow is provided by a property of the presupposition that is apparently dependent on grammatical mood. This property manifests an intuitive difference between indicative and subjunctive *because*-sentences. Based on this observation a modification of the Lewis-style formulation of the presupposition is provided for the case of counterfactual *because*.

With regards to indicative *because*-sentences the Lewisian formulation of part (ii) in the truth condition stated above in section 4.4.2 on page 146 is completely intuitive. Two examples shall serve as a reminder of what is at issue:

- (26) John got sick, because he drank too much alcohol.
 (27) John embarrassed everyone at the party, because he drank too much alcohol.

In both cases it is presupposed to be a fact that neither would have John become sick, nor would have everyone else been embarrassed by John, had he not drunk too much alcohol. Thus, when translating (26) and (27) as *because* q , p the presupposition can be given as $\neg q \Box \rightarrow \neg p$.

However, in the case of counterfactual *because* the Lewis-style formalization is too weak. It does not comprehensively capture the intuition connected to its facticity presupposition. Returning to the original example will help to clarify the point. Thus, consider once more:

- (28) Because he would have become sick, John did not drink any alcohol.

In this case the intuitive facticity presupposition is not simply:

- (29) If John would not (have) become sick, he would have drunk alcohol.

Rather – and more adequately –, the facticity presupposition gives a more elabo-

CHAPTER 4. THE ANALYSIS

rate account of the peculiar circumstances of the scenario. The presupposition is in fact:

- (30) If John would not have become sick, if he had drunk alcohol, he would have drunk alcohol.

In this the embedded *if*-clause *If he had drunk alcohol* intuitively serves two functions at once. For one it provides (as has already been observed multiple times) the necessary contextual support for the interpretation of the *would have* + Past Participle-complex given in the sentence initial *if*-clause. For the other, and intimately related to the former, the embedded *if*-clause figures as a semantic modification of the initial *if*-clause: it expresses a restriction to the given premise. And it suggests itself to conceive of the iteration of antecedents as a form of conjunction. The following paraphrase of the example underlines this intuition:

- (31) If John (could have) had drunk alcohol and not become sick, he would have drunk alcohol.³²

Pondering further upon the logical structure of this complex premise set, it is compelling to assume that when you can drink and not become sick, it holds that it is indeed not the case that if you drink you will become sick. (In this the premise actually coheres with the following tautology that clarifies the relation between ‘ \wedge ’ and ‘ \rightarrow ’: $(P \wedge \neg Q) \Leftrightarrow \neg(P \rightarrow Q)$.³³) –Which brings us to the core of the semantic relation between *because*-sentence and the elliptic counterfactual discourse, since what is negated here is the law-like connection played upon by the counterfactual

³²Since presuppositions are genuine semantic objects I will not bother about the peculiar syntactic structure exhibited in examples (30) and (31) and their formal semantic translation. However, it shall be mentioned that the kind of iterated conditional under discussion here differs significantly from those with which Huitink (2008) is concerned. While in the present example the iteration occurs in the antecedent of the (matrix)conditional (the formal structure being $((p \rightarrow q) \rightarrow r)$), Huitink discusses cases in which iteration pertains to the consequent of the matrix clause, the bracketing being here $(p \rightarrow (q \rightarrow r))$. Only the latter kind of iteration has an analogue in the *exportation law* of propositional logic: $((P \rightarrow (Q \rightarrow r)) \Leftrightarrow (P \wedge Q) \rightarrow R)$. The equivalence to $(p \wedge q) \rightarrow r$) does not hold for $((p \rightarrow q) \rightarrow r)$.

³³At least if one considers only ‘normal’ scenarios and absconds from the assumption of some extraordinary eventualities to obtain that present restraining conditions to the relation between drinking and sickness (like e.g. the pre-emptive consumption of Aspirin or another remedy against hang-overs before one starts drinking).

CHAPTER 4. THE ANALYSIS

conditional implicit in the *because*-sentence *John did not drink, because (if he had) he would have become sick*, i.e. the assumption that *If you drink too much, you will become sick*.

Brought to the point, the presupposition of the *because*-sentence in question is hence that had it not been the case that drinking and being sick were (nomologically) related (as given by $\Box(p \rightarrow q)$, with p translating someone is *drinking*, and q translating someone being *sick*), then it would have been the case that John had drunk something. In other words, the presupposition is that had the law $\Box(p \rightarrow q)$ not been established, then John would have drunk. Formally restated the modified presupposition of the *because*-paraphrase of the discourse is thus:

$$\neg\Box(p \rightarrow q) \Box\rightarrow p$$

Resp. in terms of a support test:

$$\neg\Box(p \rightarrow q) \models p$$

This is an important finding, because it completes the account of the synonymy of the *because*-data and the discourse: not only are the asserted semantic objects identical in both data, but also the presupposition of the *because*-sentence coincides trivially with the semantics of the discourse. Since for the particular example the antecedent of the presupposition is identical to the already presupposed law concerning the relation of *drinking* and *sickness*, the presupposition of the *because*-sentence involves just the information already presupposed in the discourse that it paraphrases. –Which is the last step in the account of the synonymy of the data.

4.4.4 Aside: why we all are counterfactual drunkards – the tenability of the presupposition of *because*

Although the analysis of the synonymy of the discourses and the *because*-paraphrase has been completed, one last question shall be pursued in this connection. The answer to this will turn out to be of general interest for the discussion.

CHAPTER 4. THE ANALYSIS

When pondering further upon the presupposition of the *because*-sentence doubts about its validity may arise: it is rather easy to imagine scenarios in which the subject in question will not drink even though one assumes that drinking does not result in sickness. The counterfactual expressed in the presupposition thus seems for principal reasons rather dubious. Why is the counterfactual – resp. the asserted *because*-sentence – still intuitively acceptable? What scheme of inference underlies the counterfactual?

Taking a slightly different perspective what is at issue is as follows. Given the premise that *Drinking too much alcohol results in sickness* one faces two mutually exclusive alternatives: either one drinks, or one does not drink, with the consequence of the latter (which is to not become sick) being preferred over the consequence of the former (i.e. to become sick). Hence, most people will choose not to drink. But why is it that whenever the premise that postulates the relation between sickness and drinking is let go one seems to be compelled to drinking? What logic underlies the counterfactual expressed in the presupposition?

The answer turns out to be rather natural if one conceives of the preference relation between the consequences of *drinking* respectively *not drinking* in more elementary terms. Instead of thinking about it in relational terms ('*not becoming sick* is preferable over *becoming sick*') it is more appropriate to assume that *to not become sick* figures as a first premise. Given this, and given the law that *Drinking will result in sickness* it then follows by *Modus Tollens* that one better does not drink, if one does not want to end up sick.

Assuming next counterfactually that the law that *Drinking results in sickness* is no longer valid, or, in other words, it is not the case that drinking (below translated by p) and sickness (translated by q) are nomologically related:

$$\neg\Box(p \rightarrow q)^{34}$$

³⁴Negation here is actually meta-linguistic negation. It has narrow scope, reading *If had not been the case that* $\neg\Box(p \rightarrow q)$, rather than wide scope, reading *If had not been the case that* $\Box(p \rightarrow q)$.

CHAPTER 4. THE ANALYSIS

then – *ceteris paribus* – this does not mean that one also is inclined to give up the premise that expresses the preference to not become sick. The proposition *to not become sick*, i.e. $\neg q$, remains a reasonable goal to be pursued. After all, when making a counterfactual assumption as much as possible of the already established information is to be preserved.

Hence, for the counterfactual scenario the following two premises are characteristic:

- (i) $\neg \Box(p \rightarrow q)$
- (ii) $\neg q$

Since (i) is equivalent to the assumption that you can *drink and not become sick* ($\neg(P \rightarrow Q) \Leftrightarrow (P \wedge \neg Q)$ is a propositional tautology) it follows from the definition of ‘ \wedge ’ and the assumed validity of (i) and (ii) that p (= *drinking*) must also be true. This however explains how it could be that whenever one assumes that the law that *Drinking too much alcohol leads to sickness* were false, one automatically is led to assume that people would drink. Which of course renders the counterfactual support test at the heart of the presupposition of the *because*-sentence intuitively successful.

Chapter 5

Conclusion

This concluding section will combine two things: for one I will deliver a brief summary of what has been discussed and achieved in this thesis, and for the second, an outlook on what topics could (or should) further be studied shall be given.

Being generally concerned with the semantics and pragmatics of counterfactual conditionals, the initial point of departure for the discussion was given by an asymmetry in discourse felicity of a certain kind of elliptic counterfactual discourse. These discourses consist of a first sentence which gives some factual information which is next followed by the statement of what would have been the case if this initially mentioned (positive or negative) fact had not obtained. Or, more precisely, the discourse initial sentence $\pi(\varphi)$ (= sentence S1) explicitly expresses the presupposition π connected to the implicit antecedent φ of the following elliptic counterfactual conditional $\varphi \square \rightarrow \psi$ (= sentence S2):

$$[\pi(\varphi)]_{S1} \cdot [\varphi \square \rightarrow \psi]_{S2}$$

The observation connected to this kind of data was that it is not always felicitous to leave the antecedent of the counterfactual implicit in such a sequence. Indeed, the phenomenon seems to be systematic: the opposite of an acceptable sequence appears to be invariantly infelicitous. I.e. whenever you are able to give a fact p and felicitously continue with what would have been the case if this fact had not obtained, the variant of this discourse which is concerned with the counterfactual

CHAPTER 5. CONCLUSION

consequence of the opposite scenario (which ultimately states that since $\neg p$, the consequence of p could not have been) turns out to be not an acceptable discourse.

This strange behavior of the data – which I dubbed *Veltman's puzzle* due to it being observed first in Veltman (2005) – pressed for answers to the following questions: What are the felicity conditions for elliptic counterfactuals of this kind? When may the antecedent of a counterfactual conditional be left implicit in discourse? And, intimately related to this, why do we understand such elliptic discourses in the way we do understand them? What determines the implicit antecedent?

Since the initial set of data suggested that negation might be a – if not *the* – crucial factor for the felicity of the discourses (in the first set of examples examined the felicitous discourse opens with a negated sentence, while the first sentence of the infelicitous discourse does not contain a negation) the analysis focussed on the role and contribution of negation to phenomenon. This eventually led to the identification of what I call *negation-induced alternatives* (NinAs). NinAs represent implicit, contextual (indeed counterfactual) elements. Confirming the factual availability of the NinAs by observing the synonymy of discourses like *John did not drink. He would have become sick* and *John did not drink. Otherwise he would have become sick* in which the particle *otherwise* figures as a semantic discourse mirror seemed to support the initial idea that discourse felicity is determined by the availability of a NinA (which could act as an implicit filler for the hidden antecedent). However, the fact that there were examples that did not fit the picture, but contradicted the account was rather disturbing.

Being thus forced to reconsider the problem, a turn of attention to the *because*-sentence paraphrase of the discourses under consideration opens up the chance to get a new perspective on the issue. This change in perspective eventually provides with a solution to Veltman's puzzle that is compact, general, and in accordance with the semantic essence of conditional statements, which after all relate conclusions with their premises. It takes its departure with the following observation:

Whenever intuition is followed and (counterfactual) conditionals are understood

CHAPTER 5. CONCLUSION

to relate conditions with the consequences deriving from them, then in order for the consequent to be true at least one of the conditions sufficient for the consequent has to be true. Somewhat more formally this can be captured as follows:

$$p_1 \vee p_2 \vee p_3 \dots \Box \rightarrow q$$

Or, put somewhat differently, at least one of the possible premises must support the conclusion:

$$p_1 \vee p_2 \vee p_3 \dots \models q$$

In this case the premises thus can be conceived to build a *disjunctive* array of possible antecedents.

Turning to the opposite it holds that for the consequent not to be true, it is necessary that not a single of the conditions entailing (resp. supporting) it is true. This yields accordingly a *conjunctive* premise (set):

$$\neg p_1 \wedge \neg p_2 \wedge \neg p_3 \dots \models \neg q$$

The logic that characterizes the relation between disjunctive and conjunctive premise sets is given by the following equivalence:

$$(p_1 \vee p_2 \vee p_3 \vee \dots p_n) \leftrightarrow q \quad \Leftrightarrow \quad (\neg p_1 \wedge \neg p_2 \wedge \neg p_3 \wedge \dots p_n) \leftrightarrow \neg q$$

Applying these rather general considerations to the case of elliptic counterfactuals the account of the infelicity of an *unacceptable* discourse is actually a matter of semantic evaluation rendered impossible. With the antecedent remaining implicit, the space of what could possibly be assumed remains open. (It arguably is indeed infinite, since assumptions are essentially free.) This however means that if it is necessary for the consequent to be true that all sufficient conditions are to be excluded – i.e. whenever the set of premises takes the form of a conjunction – the validity of the consequent can not be evaluated: while in case an explicit antecedent had been given, one could have concluded by way of a quantity implicature that all other possible antecedents than the one mentioned are false and thus need not be checked, no such conclusion is possible in case of an elliptic

CHAPTER 5. CONCLUSION

conditional with implicit antecedent. Here in principle all possible antecedents would have to be considered in order to determine whether the consequent is true or not. This however is not possible, because – as has already been noted – the set of possible premises is unrestricted and indeterminate. Hence, the discourse becomes infelicitous. The elliptic sequence is not interpretable, because there is no way to evaluate the counterfactual consequent.¹

This summarizes in a nutshell the solution to Veltman’s puzzle. This solution is the first main result of the thesis. Another result that has been formulated is concerned with what I dubbed the ‘radical interpretation’ of counterfactual conditionals. It is connected to conditionals like *If John had not drunk, he would not have become sick* that lack any sound empirical or nomological support, and that are – as we have seen – actually semantically false. Nonetheless unsupported conditionals of this kind are usually still accepted as true in conversation. How this could be justified is a point that to my knowledge has not been addressed in the literature. Pursuing this issue, two lines of arguments are discussed. The more elaborate line of reasoning involves the concept of conditional strengthening, for which two accounts are given: the first one involves the adaptation of an illocutionary operator that was formulated in Krifka (1999) (cf. section 1.7.1), while the other (which is more problematic², but in its gist closer to some of the ideas that gave rise to the solution of Veltman’s puzzle) employs von Stechow’s [2001] account of conditional strengthening (cf. section 4.2.4, p. 134 ff.). Simpler in its dialectic and an alternative to the conditional strengthening-approaches is the line of reasoning elaborated in section 4.2.2. While also in this account a quantity implicature plays a crucial role, its argumentative force rests on more general assumptions about rationality than the competing accounts in terms of conditional strengthening. What is common to all these approaches is however that they allow for a more fine grained account and deeper understanding of the interplay between background knowledge and the semantic and pragmatic resources involved in the interpretation of counterfactual conditionals.

¹The reason why indicative conditionals do not behave in the same way is that only in the subjunctive case can a consequent of a conditional actually be identified as such. There are no elliptic conditional discourses in the indicative.

²A major drawback of this account is the problem pointed out in footnote 21, on page 136.

CHAPTER 5. CONCLUSION

While also in connection with the *because*-paraphrase of the discourses under discussion a number of results have been achieved, at least some other points that have so far not been addressed in detail shall now also be mentioned. The first of these points concerns a tacit assumption on which the analysis is based. The account presupposes that the interpreter of an elliptic counterfactual has material knowledge about event structure and possible relations between events. Language alone does not determine what premises might figure as sufficient conditions for which consequent. Knowledge of this kind however is crucial to differentiate a disjunctive from a conjunctive premise set, which in turn decides discourse felicity.³ Once this material aspect underlying interpretation is realized a bunch of further intricacies of the solution become apparent. The following data exemplify at least some of these. The first one to be mentioned highlights an issue intimately related to the question whether negation in the consequent invariably determines the premise set to be conjunctive, resp. whether the logical structure of the premise set can reliably be determined via the question under discussion. That this is not always clear is demonstrated for instance by the verb *to fall asleep*.⁴ Pondering the question

When do you fall asleep?

the answer may be disjunctively given:

If you are tired, OR if the movie is boring, OR if you took sleeping pills, OR if...

Considering the negation to this case one might expect – as has been just expounded – that the answer delivers a conjunctive set of possible antecedents. And although the absence of all the premises just given describes a necessary condition for *not* to fall asleep, one might however also think of conditions that build a disjunctive premise set that answers the question

³The problem of how to identify the ‘relevant’ conditions for the truth of a counterfactual consequent is already addressed in e.g. Goodman (1947). That material knowledge plays an important role actually seems to support the stance taken in van Lambalgen and Hamm (2005) who consider the capacity to interpret natural language to be an exaptation of the human planning faculty and our understanding of causality. Following this it would be highly interesting to formulate a hybrid semantics of Veltman’s theory of counterfactuals and the formalism developed in van Lambalgen and Hamm (2005).

⁴The example is due to Jerra Lui Busch, p.c.

CHAPTER 5. CONCLUSION

When do you not fall asleep?

These are for instance:

If you are not tired, OR you drink enough coffee, OR if it is too noisy, OR if . . .

These cases also provide sufficient conditions that would render the conclusion of a counterfactual true. And indeed, the following two discourses are felicitous:

- (1) John did not drink coffee. He would not have fallen asleep.
- (2) John drank coffee. He would have fallen asleep.

While the introduction of negation into the consequent in case of the examples considered so far (like the *drinking/becoming sick*-scenario discussed in the previous chapters) immediately lead to the differentiation of the logical structure of the premise set as either conjunctive or disjunctive, in the above example negation does not enforce this distinction. Apparently, in both cases the default for evaluation is the disjunctive premise set. What triggers this behavior of the data is not clear to me and needs to be studied further. It might be ontological or psychological factors, and it would be in any case interesting to have an answer to this question.

The fact that there is a disjunctive default is arguably also the reason, why the corresponding opposites in (3) and (4) are at least in my intuition not completely infelicitous:

- (3) ?John did not drink coffee. He would have fallen asleep.
- (4) ?John drank coffee. He would not have fallen asleep.

With the premise set of the consequent being disjunctively structured, the discourses arguably facilitate to accommodate a rule-like fact that (in combination with the given factual (and presuppositional) information) supports the conclusion. In case of (3) the discourse is rendered felicitous if one assumes for instance a scenario in which the coffee is poisoned with a high dose of barbiturates. Thus accommodated, the discourse is perfectly felicitous and the consequent true, since

CHAPTER 5. CONCLUSION

although drinking coffee ‘normally’ prevents falling asleep this will not be the case if the coffee contains enough sleeping pills. In this case one might drink coffee, but still fall asleep. A similar accommodation of the coffee containing sleeping pills guarantees also a felicitous reading of (4). Based on examples like these the mechanism of accommodation becomes just the more interesting. Lewis cautious ‘definition’ leaves enough room for further, more detailed study from both, the theoretical, and the empirical perspective. That the non-monotonicity of counterfactuals plays an important role at this point is obvious. Indeed, the question whether the possibility to accommodate is based on the non-monotonicity of counterfactuals or their non-monotonicity is due to the possibility to accommodate is a question that is highly interesting and should be studied further. It actually touches upon fundamental problems concerning the nature of rules (and their relation to counterfactuals) that find discussion for instance in [Goodman \(1947\)](#) but that still remain open.

Accommodation is central also to another point that has not been addressed in any detail. The observation is that modals like *must* and *can* apparently exert a certain influence on the acceptability of the data. Consider for instance the following example that was already mentioned in footnote 22 on p. 23:

(5) John did not win the race. He would have celebrated.

For this discourse the *because*-paraphrase is intuitively marginal:

(6) ?Because he would have celebrated, John did not win the race.

However, the picture changes when a deontic *to have to* – that like *must* comes with a universal modal force – is inserted into the second sentence:

(7) John did not win the race. He would have had to celebrate.

For this case the *because*-paraphrase is felicitous, determining a reading that John intentionally did not win the race in order to not have to celebrate:

(8) Because he would have had to celebrate, John did not win the race.

CHAPTER 5. CONCLUSION

The modal facilitates the accommodation of a motive that could explain the subject's behavior (like for instance that John did not win the race because he – although it is customary to celebrate winners – wanted to avoid having to celebrate, since he absolutely hates being at the center of attention).

And similar for the insertion of the modal *can* into the discourse initial sentence:

(9) John can not have won the race. He would have celebrated.

Here the modal enforces an evidential reading that asks the hearer pretty much to *take a look around and note that no signs of a celebration can be found*. In all these cases there is a clear interplay between the occurring modal, the expressed discourse relation and what can be accommodated.

Another related question (but this time connected to the nature of the premise set) is what happens in case the set of possible premises is clearly delineated and well defined. If there is a case which turns out true on just a single condition not only would it be impossible to differentiate the premise set as either disjunctive or conjunctive, but since the set of possible antecedents is finite this might affect acceptability judgements. And apparently – at least at first sight – there are such cases. Just consider a scenario in which a soup is *oversalted*⁵. For this there exists (naively) just a single 'explanation': a soup is oversalted, if and only if too much salt is put into it. No other 'cause' seems possible. However, despite this, my first intuition is that certain examples are acceptable, while others are not:

(10) John cooked the soup. It would have been oversalted.

(11) John did not cook the soup. It would have been oversalted.

(12) # John cooked the soup. It would not have been oversalted.

(13) # John did not cook the soup. It would not have been oversalted.

Apparently, while there is indeed only a single 'direct' cause imaginable for the soup being too salty, it is possible to accommodate the scenario in such a way

⁵The example is again due to Jerra Lui Busch, p.c.

CHAPTER 5. CONCLUSION

that alternative premises become nonetheless available. The alternative scenario in case of (10) is one in which other people might have cooked the soup (and all of them would have put too much salt into it), while case of (11) the existence of other agents might not be accommodated, and the discourse is ambiguous with respect to whether the soup is cooked at all. Apparently, in (12) it is not possible to exclude that no one else besides John would not have put too much salt into the soup, while in case of (13), presupposing that the soup has actually been cooked, we do not want to universally deny the possibility that John could have oversalted the soup. Cursory as these considerations are, they should give already an idea what questions and directions might further be pursued.

Bibliography

- Anderson, Alan Ross. 1951. A Note on Subjunctive and Counterfactual Conditionals. *Analysis* 12:35 – 38.
- Baker, Carl LeRoy. 1966. *Definiteness and indefiniteness in English*. Master's thesis, University of Illinois.
- Biezma, María, and Kyle Rawlins. 2011. Responding to Alternative and Polar Questions. DRAFT, accessed 04.09.2012, 23:37h, under <http://semanticsarchive.net/Archive/2JmZGE5N/Responding>
- Carlson, Greg N. 1977. Amount Relatives. *Language* 53:520 – 542.
- Cornulier, Benoît de. 1983. 'If' and the Presumption of Exhaustivity. *Journal of Pragmatics* 7:247 – 249.
- Davidson, Donald. 1973. Radical Interpretation. *Dialectica* 27:314 – 328.
- Dowty, David R. 1979. *Word Meaning and Montague Grammar*. D. Reidel Publishing Company, Dordrecht.
- von Fintel, Kai. 2001. Conditional Strengthening: A Case Study in Implicature. Unpublished draft, version from 9/13/2001, 1:08 PM.
- Frank, Anette. 1997. *Context Dependence in Modal Constructions*. Doctoral Dissertation, Universität Stuttgart. Published in: Arbeitspapiere des Sonderforschungsbereichs 340, 'Sprachtheoretische Grundlagen für die Computerlinguistik', Nr. 91.

BIBLIOGRAPHY

- Frank, Anette, and Hans Kamp. 1997. On Context Dependence in Modal Constructions. In *Proceedings of the 7th Semantics and Linguistic Theory Conference (SALT 7), held March 21-23, 1995, at Stanford University*, ed. Aaron Lawson, 151–168. Ithaca NY: Cornell University.
- Fuhrmann, Katharina. 2011. *The Meaning of 'otherwise'*. Master's thesis, Georg-August-Universität Göttingen.
- Geach, Peter. 1967. Intentional Identity. *The Journal of Philosophy* 64:p. 627 – 632.
- Geis, Michael L., and Arnold M. Zwicky. 1971. On Invited Inferences. *Linguistic Inquiry* 2:561 – 566.
- Ginzburg, Jonathan. 2012. *The Interactive Stance: Meaning for Conversation*. Oxford University Press, Oxford.
- Goodman, Nelson. 1947. The Problem of Counterfactual Conditionals. *The Journal of Philosophy* 44:113 – 128.
- Grice, Herbert Paul. 1975. Logic and Conversation. In *Speech Acts*, ed. Peter Cole and Jerry L. Morgan, volume 3 of *Syntax and Semantics*, 41–58. New York: Academic.
- Groenendijk, Jeroen, and Martin Stokhof. 1984. *Studies on the Semantics of Questions and the Pragmatics of Answers*. Doctoral Dissertation, Universiteit van Amsterdam.
- Groenendijk, Jeroen, and Martin Stokhof. 1991. Dynamic Predicate Logic. *Linguistics and Philosophy* 14:39 – 100.
- Heim, Irene. 1982. *The Semantics of Definite and Indefinite Noun Phrases*. Doctoral Dissertation, University of Massachusetts at Amherst.
- Heim, Irene. 1992. Presupposition Projection and the Semantics of Attitude Verbs. *Journal of Semantics* 9:183–221.

BIBLIOGRAPHY

- Hsieh, I-Ta Chris. 2009. On NPI Licensing and the Semantics of Causal Sentences. In *Proceedings of the 10th Symposium on Logic and Language*, ed. László Kálmán, 63 – 72. Budapest: Research Institute for Linguistics, Hungarian Academy of Sciences (HAS), and Theoretical Linguistics Program, Eötvös Loránd University (ELTE).
- Huitink, Janneke. 2008. Belnap-style Conditional Semantics. Hand-out of a talk given at the Department of English Language and Literature, Georg-August-Universität Göttingen, June 2, 2008.
- Hume, David. 1999. *An Enquiry Concerning Human Understanding*. Oxford University Press, Oxford. First published: 1748.
- Jackendoff, Ray S. 1972. *Semantic Interpretation in Generative Grammar*. MIT Press, Cambridge, MA.
- Kamp, Hans. 1981. A theory of truth and semantic representation. In *Formal Methods in the Study of Language: Mathematical Centre Tracts, Nr. 135* (= *Proceedings of the third Amsterdam Colloquium on Formal Methods in the Study of Language, held March 25-28, 1980*), ed. Jeroen Groenendijk, Theo Janssen, and Martin Stokhof, 277–322. Mathematisch Centrum, Amsterdam.
- Kamp, Hans, and Uwe Reyle. 1993. *From Discourse to Logic: Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*. Kluwer, Dordrecht.
- Karttunen, Lauri. 1974. Presupposition and Linguistic Context. *Theoretical Linguistics* 1:181 – 194.
- Kasper, Walter. 1987. *Semantik des Konjunktivs in Deklarativsätzen des Deutschen*. Max Niemeyer Verlag, Tübingen.
- Kaufmann, Stefan. 2000. Dynamic Discourse Management. In *Formalizing the Dynamics of Information*, ed. Martina Faller, Marc Pauly, and Stefan Kaufmann, 171–188. CSLI Publications, Palo Alto, CA.

BIBLIOGRAPHY

- Kaup, Barbara, Jana Lüdtkke, and Rolf Zwaan. 2006. Processing negated sentences with contradictory predicates: is a door that is not open mentally closed? *Journal of Pragmatics* 38:1033 – 1050.
- Kaup, Barbara, Richard Yaxley, Carol Madden, Rolf Zwaan, and Jana Lüdtkke. 2007. Experiential Simulations of Negated Text Information. *The Quarterly Journal of Experimental Psychology* 60(7):976–990.
- Kratzer, Angelika. 1989. An Investigation of the Lumps of Thought. *Linguistics and Philosophy* 12:607 – 653.
- Krifka, Manfred. 1999. At Least Some Determiners Aren't Determiners. In *The Semantics/Pragmatics Interface from Different Points of View (= Current Research in the Semantics/Pragmatics Interface Vol. 1)*, ed. Ken Turner, 257 – 291. Elsevier Science, Amsterdam.
- van Kupperfelt, Jan. 1996. Inferring from Topics. *Linguistics and Philosophy* 19:393 – 443.
- Lakoff, George. 1970. Linguistics and natural logic. *Synthese* 22:151 – 271.
- van Lambalgen, Michiel, and Fritz Hamm. 2005. *The Proper Treatment of Events*. Wiley-Blackwell, Malden, MA.
- Lewis, David K. 1973a. Causation. *Journal of Philosophy* 70:556 – 567.
- Lewis, David K. 1973b. *Counterfactuals*. Basil Blackwell, Oxford.
- Lewis, David K. 1979a. Counterfactual Dependence and Time's Arrow. *Noûs* 13:455 – 476.
- Lewis, David K. 1979b. Scorekeeping in a Language Game. *Journal of Philosophical Logic* 8:339 – 359.
- Matsumoto, Yo. 1995. The Conversational Condition on Horn Scales. *Linguistics and Philosophy* 18:21 – 60.

BIBLIOGRAPHY

- McCawley, James D. 1981. *Everything that Linguists Have Always Wanted to Know About Logic (but were Ashamed to Ask)*. University of Chicago Press, Chicago, IL.
- Montague, Richard. 1973. The Proper Treatment of Quantification in Ordinary English. In *Approaches to Natural Language*, ed. Jaakko Hintikka, Julius Moravcsik, and Patrick Suppes. Reidel, Dordrecht.
- Pollock, John L. 1976. *Subjunctive Reasoning*. Reidel, Dordrecht.
- Ramsey, Frank P. 1929. General Propositions and Causality. In *Philosophical Papers*, ed. David H. Mellor, 145–163. Cambridge University Press, Cambridge, MA (1990).
- Reiter, Ray. 1978. On Closed World Data Bases. In *Logic and Data Bases*, ed. Hervé Gallaire and Jack Minker, 55 – 76. Plenum Press, New York and London.
- Reiter, Ray. 1988. On Integrity Constraints. In *TARK '88 Proceedings of the 2nd conference on Theoretical aspects of reasoning about knowledge*, ed. Moshe Vardi, 97 – 111. Morgan Kaufmann Publishers Inc. San Francisco, CA.
- Roberts, Craige. 1987. *Modal Subordination, Anaphora, and Distributivity*. Doctoral Dissertation, University of Massachusetts at Amherst.
- Roberts, Craige. 1989. Modal Subordination and Pronominal Anaphora in Discourse. *Linguistics* 12:683–721.
- Roberts, Craige. 1996a. Anaphora in Intensional Contexts. In *The Handbook of Contemporary Semantic Theory*, ed. Shalom Lappin, 215–246. Blackwell, Oxford.
- Roberts, Craige. 1996b. Information Structure in Discourse: Towards an Integrated Formal Theory of Pragmatics. In *OSU Working Papers in Linguistics, Vol 49: Papers in Semantics*, ed. Jae-Hak Yoon and Andreas Kathol, 91 – 136.
- Romero, Maribel, and Chung-Hye Han. 2003. Focus, Ellipsis and the Semantics of Alternative Questions. In *Empirical Issues in Formal Syntax and Semantics*

BIBLIOGRAPHY

- 4, ed. Claire Beysade, Olivier Bonami, Patricia Cabredo Hofherr, and Francis Corblin, 291 – 307. Presses Universitaires de Paris-Sorbonne, Paris.
- Rooth, Mats. 1985. *Association with Focus*. Doctoral Dissertation, University of Massachusetts at Amherst.
- van Rooy, Robert. 2005. A Modal Analysis of Presupposition and Modal Subordination. *Journal of Semantics* 22:281 – 305.
- Schueler, David. 2008. *The Syntax and Semantics of Implicit Conditionals: Filling In the Antecedent*. Doctoral Dissertation, University of California, Los Angeles.
- Stalnaker, Robert. 1968. A Theory of Conditionals. In *Studies in Logical Theory*, ed. Nicholas Rescher, 98–112. Basil Blackwell, Oxford.
- Tian, Ye, Richard Breheny, and Heather Ferguson. 2010. Why we simulate negated information: A dynamic pragmatic approach. *The Quarterly Journal of Experimental Psychology* 63:2305 – 2312.
- Tichý, Pavel. 1976. A Counterexample to the Stalnaker-Lewis Analysis of Counterfactuals. *Philosophical Studies* 29:271–273.
- Veltman, Frank. 2005. Making Counterfactual Assumptions. *Journal of Semantics* 22:159–180.
- Wilkinson, Karina J. 1991. *Studies in the semantics of generic noun phrases: a dissertation*. Graduate Linguistic Student Association, Department of Linguistics, South College, University of Massachusetts, Amherst, MA.
- von Wright, Georg Henrik. 1963. *Norm and Action: a Logical Enquiry*. Humanities Press, New York, NY.
- von Wright, Georg Henrik. 1968. *An essay in deontic logic and the general theory of action*. North-Holland Publ., Amsterdam.