

Chisholm's Paradox in *Should*-Conditionals

Ana Arregui
University of Ottawa

1. Introduction

This paper will be concerned with the problem of factual detachment in deontic conditionals. The goal is to investigate a semantics for deontic modals and conditionals that tracks facts closely, validates factual detachment and avoids paradoxes. Illustrations will be provided making use of the deontic modal *should*. In the paper, I will simplify matters and set aside important differences between deontic modals. Further research will be needed for a more general picture.¹

2. Chisholm's Paradox

In a 1963 paper, Chisholm presented a puzzle (*Chisholm's Paradox*) to argue that the tools provided by standard deontic logic (SDL) were too limited to give a proper account of natural language deontic statements. Chisholm put together a set of English sentences that were intuitively consistent and independent of each other. He argued that the sentences could not be given a consistent and independent representation within SDL. The sentences are given in (1):

- (1) a. It ought to be that a certain man go to help his neighbours.
- b. It ought to be that if he goes, he tell them that he is coming.
- c. If he does not go, he ought not to tell them that he is coming.
- d. He does not go.

Chisholm provided the following SDL-representation for the sentences in (1) (where O is the deontic necessity operator, \rightarrow is the material conditional and abbreviations are used for the propositions): (1a) $O \text{ help}$, (1b) $O (\text{help} \rightarrow \text{tell})$, (1c) $\neg \text{help} \rightarrow O \neg \text{tell}$, and (1d) $\neg \text{tell}$. As Chisholm pointed out, this formalization results in inconsistency: from (1a) and (1b) we can derive $O \text{ tell}$, and from (1c) and (1d) we can derive $O \neg \text{tell}$ and the conjunction is a contradiction. Carmo and Jones (2002) discuss alternative formalizations, and point to the problems that arise with those: if both conditionals (1b, c) are represented with the material conditional and O in the consequent ($\rightarrow O$), we then lose independence: (1b) would be represented as $\text{help} \rightarrow O \text{ tell}$, which follows from (1d). If both conditionals (1b, c) are represented with the necessity operator followed by the material conditional ($O \rightarrow$), we also lose independence: (1c) would be represented

¹Readers are referred to Copley (2006), von Stechow and Iatridou (2006) and references cited there for a discussion of differences between modals and their quantificational strength.

by O ($-help \rightarrow tell$), which follows from (1a). The conclusion reached is that if we cannot provide a consistent independent representation of (1a-d) within SDL, we need more resources.

The problems posed by (1) can be thought of as a *detachment* puzzle. We will use the word *detachment* as a cover term for a general inference pattern: from a conditional sentence of the form *if* α , β and the premise α , we are able to infer β . Detachment has been discussed in two variants: *factual detachment* and *deontic detachment*. In the first case, detachment is justified by the truth of α , in the second, detachment is justified by the truth of *should* α . If we represent the conditionals in (1) with the toolkit provided by SDL, we predict that both types of detachment will be valid. Indeed, both types of detachment options play a role in drawing out the SDL-contradiction in (1): we see factual detachment between (1c, d) and deontic detachment between (1a, b). The puzzle presented by the sentences in (1) can be thought of as a detachment puzzle because with SDL both detachment options are validated and given the semantics of O , the result is a contradiction.

To see the intuitions associated with the different detachment options, consider the examples below:

- (2)
- a. You: If she returns her library book late, she should pay a fine.
 - b. Me: She returned her library book late (*uttered some days later*)
 - c. You: Aha! She should pay a fine.

In (2) I have presented a little dialogue set in time (sometime between your utterance and mine, she returns the book late). Looking at the little example, we have the intuition that (2c) is true. In order to claim that (2c) follows from (2a, b) it is necessary to provide a semantics for the *should*-conditional that validates factual detachment. (Note: In this paper, I will try to make sense of the idea that (2c) does indeed follow from (2a, b), though, as we will see, this is not simple. However, there could be other reasons that make us believe that (2a) and (2b) together give rise to an unconditional obligation. Maybe there is pragmatic pressure to accept (2c) given (2a, b), even if (2c) does not really follow from (2a, b). However, a theory predicting this has not been worked out, and I will set this option aside here.)

The intuitions regarding deontic detachment with *should*-conditionals are less straightforward. Consider the variant of the example in (3):

- (3)
- a. You: She should return the library book on time.
 - b. Me: If she returns the library book on time, she should be allowed to borrow an extra book next time she comes to the library.
 - c. You: Well, she should be allowed to borrow an extra book next time she comes to the library.

There is something not quite right about your reasoning in (3). You seem rather hasty. The truth of the deontic statement that she should return the library book on time does not move us immediately to the conclusion in (3c). Yet, this would be

predicted if deontic detachment were valid for *should*-conditionals. I mention deontic detachment here for completeness, but will set it aside in this paper.

Much of the discussion surrounding Chisholm's Paradox has focused on examples like (1c). This is a conditional that tells us what should happen if less-than-ideal circumstances hold (if the man does not go to help his neighbours). Deontic conditionals that tell us what to do in less-than-ideal circumstances have been dubbed 'Contrary To Duty Imperatives' (CTDs). They are important, according to Chisholm, *For most of us need a way of deciding, not only what we ought to do, but also what we ought to do after we fail to do some of the things we ought to do* (Chisholm 1963: 35-36). CTD conditionals have received much attention in the logical and philosophical literature on deontic conditionals (a.o. Åqvist 2002, Carmo & Jones 2002.). Factual detachment operating on these conditionals leads us to conclusions that get us into trouble.

3. Dealing with Deontic Conditionals

Proposals to deal with deontic modality have often moved away from SDL. In this section we will discuss a dyadic operator approach following the work of Lewis (1973) and Feldman (1986). This is a proposal to deal with deontic conditionals by means of a 'modal' conditional operator that takes two arguments (instead of the material conditional we see in SDL). The approach does not validate factual detachment, and so does not give rise to inconsistencies with CTDs. The proposal does validate deontic detachment (for an informal proof the reader is referred to Åqvist 2002). The purpose of this section is to provide us with some background to better evaluate the proposal that will be presented in Section 5.

3.1 Lewis (1973)

In his book *Counterfactuals* (Lewis 1973), Lewis proposed a 'variably strict' analysis of counterfactual conditionals in terms of a dyadic operator that combines with two propositions. The semantics of the conditional operator was stated in terms of a system of nested spheres that encoded relative similarity. In his book, Lewis also proposed to extend the structure of nested spheres to provide a semantics for deontic conditionals. The Lewis-semantics for counterfactuals claimed that a counterfactual was true (given a system of nested spheres) iff the consequent was true in the worlds of the sphere in which the antecedent was true. This meant that for the counterfactual to be true, the consequent had to be true in the most similar worlds in which the antecedent was true (simplifying a little and adopting the Limit Assumption). In Lewis's deontic reinterpretation, the construction of the system of spheres corresponds to the amount of goodness in a world (for whatever measure is relevant). Worlds in the same sphere count as 'equally good', with the relative ordering within the system of spheres corresponding to the relative goodness of the worlds (worlds in the closer spheres are better than worlds in the further spheres). A deontic conditional is true (given a system of spheres) if the consequent is true in the worlds of the sphere in which

the antecedent is true. This means that for a deontic conditional to be true, the consequent has to be true in the best worlds in which the antecedent is true.

Lewis's proposal for deontic conditionals does not validate factual detachment. With the Lewis-semantics, a deontic conditional claims that in the best worlds in which the antecedent is true, the consequent is also true. Truth of the antecedent in the actual world will give rise to no inferences. The actual world need not be one of the best worlds in which the antecedent is true, so nothing follows from the mere actual truth of the antecedent. Since Lewis's system does not validate factual detachment, it can be used to provide a consistent and independent representation for Chisholm's sentences. However, with this proposal we would lose our intuitions regarding (2).

3.2 *Feldman (1986)*

Lewis's proposal for deontic conditionals abstracts away from temporal matters. Yet, intuitively, our judgements regarding Chisholm's sentences in (1) seem affected by our temporal perspective (though Chisholm did not discuss this). Intuitions tell us that claims regarding what is best can vary depending on when they are made. *He ought to tell the neighbours that he is going to help*, for example, is judged true at a time when his helping is a 'live option', and not afterwards. In this section we will present (parts of) Feldman's account of deontic modality (Feldman 1986). This account will be of particular interest to us because it builds on Lewis's work, but makes explicit the role of times.

Our judgments regarding what is best seem to change through time, and one way to explain this is to propose that what is best is established relative to part of the history of the world (maybe the history of the world up to a certain time). Temporal variations in our judgements can be explained as variations with respect to the pieces of the history of the world that are being taken into account. What is best relative to a particular stretch may be different from what is best relative to a larger stretch.

Here we will discuss a general notion of deontic necessity that Feldman terms 'ought-to-be' (OB). Feldman (as Lewis) set up a system that compared the relative goodness of worlds. He made use of a function IV that delivered the intrinsic value of each world (for whatever measure of value was relevant). Feldman differed from Lewis in being explicit about the temporal anchoring of deontic necessity. In figuring out what ought to be at a time t , we consider all the worlds that are still physically possible at t , and see what happens in the best of them. If p occurs in one, and $\neg p$ doesn't occur in any as good or better, then at t it ought to be that p . The world accessibility relation (WA) is roughly characterized as follows: for all worlds w , w' and times t , $WA w', w', t$ if as of t , w' has not been ruled out by what has physically happened in w . The relation of world accessibility (WA) is temporally sensitive. As time goes by, the set of worlds accessible from the actual world becomes smaller. Less and less is possible. With this view of accessibility at hand, we can define what ought-to-be (OB) as follows:

- (4) OB t, p is true at w iff $\exists w'$ (WA w', w, t & p is true at w' & $\neg \exists w''$ (WA w'', w, t & $\neg p$ is true at w'' & $IV(w'') \geq IV(w')$)) (Feldman 1986: 186)

Given (4), it ought to be that p at a world w at a time t iff p is true in a world physically accessible from w at t and there isn't a world accessible from w at t in which p is false that is as good or better.

Having a notion of deontic necessity with a temporal anchor, it is possible to define a dyadic conditional operator (q/p) that is also sensitive to times:

- (5) OB $t, q/p$ is true at w iff $\exists w'$ (WA w', w, t & p and q are true at w' & $\neg \exists w''$ (WA w'', w, t & p and $\neg q$ are true at w'' & $IV(w'') \geq IV(w')$)) (where q/p should be read 'q if p')

Like Lewis's dyadic operator, Feldman's dyadic operator fails to validate factual detachment and does validate deontic detachment. To see an example of Feldman's proposal at work, consider the following scenario and discussion (adapted from Feldman 1986): A doctor is in charge of administering medication to a patient. The best option would be to give medicine A on Monday and Tuesday (this is what the doctor should do). However, if the doctor gives medicine B on Monday, she should also give medicine B on Tuesday (all other courses of action would be fatal). Imagine now the following sentences, uttered on Saturday (before the medicine is to be administered):

- (6) a. If the doctor gives the patient B on Monday, she should also give him B on Tuesday.
 b. The doctor will give the patient B on Monday.
 c. The doctor should give the patient A on Monday and A on Tuesday.

On Saturday, all three sentences may well be true. And Feldman's proposal allows for this. On Saturday, when the doctor can still choose her course of action regarding Monday (and thus an A-on-Monday world is accessible), the doctor has the obligation of giving the patient A on Monday and A on Tuesday. The fact that she will (maybe by mistake, maybe on purpose) give the patient B on Monday is not enough for us to accept that on Saturday she has the obligation to give the patient B on Tuesday (within the set of worlds accessible on Saturday, there is a world in which the doctor gives A on Monday and A on Tuesday and there isn't a world as good or better in which she does not give A on Tuesday). With Feldman's proposal, we will not be able to put together (6a, b) to derive that the doctor should give the patient B on Tuesday (and thus there are no contradictions with (6c)).

4. *Should vs. Should Have*

In this section we will investigate the contrast between *should* and *should have*. This will help set the stage for the discussion in Section 5.

In Section 3 we presented an account of deontic conditionals that does not validate factual detachment and thus avoids contradictions in dealing with Chisholm's sentences. Without factual detachment, we will not reach the conclusion in (1) that the man ought to help his neighbours and he ought to tell them he will not go to help. But we are left with the intuition that he ought to have gone to help them. Even if the moment for helping is past, and we would not want to say that he should help them, we might still want to say that he should have helped them. This section will be dedicated to the difference between *should* and *should have*. We will examine the idea that the difference is simply a temporal difference. And we will argue that it is more than that.

4.1 A Temporal Perspective

We will take as a starting point Feldman's proposal in (4), which is interesting to us here because it makes explicit the role of a temporal parameter in establishing world accessibility (for other proposals that have relativized world accessibility to a temporal parameter in a different context, see a.o. Ippolito 2003). Having an explicit temporal parameter, it would be possible to characterize *should have* as a past form of *should* (and thus reduce the contrast between the two forms to a temporal contrast). An implementation of this idea could look as follows:

- (7) $\llbracket \text{should} \rrbracket =$
 $\lambda p \lambda t \lambda w \exists w' (WA w', w, t \ \& \ p(w') = 1$
 $\ \& \ \neg \exists w'' (WA w'', w, t \ \& \ IV(w'') \geq IV(w') \ \& \ p(w'') = 0))$
- (8) $\llbracket \text{have} \rrbracket =$
 $\lambda P \lambda w \exists t (t \prec_{\text{precedes}} s^* \ \& \ P(t)(w) = 1)$, where s^* is the speech time in the context and P is a variable ranging over properties of times.

With (7-8), a *should have p* statement will have the LF and denotation in (9):

- (9) a. $\llbracket \text{have} [\text{should } p] \rrbracket$
 b. $\lambda w \exists t (t \prec_{\text{precedes}} s^* \ \& \ \exists w' (WA w', w, t \ \& \ p(w') = 1$
 $\ \& \ \neg \exists w'' (WA w'', w, t \ \& \ IV(w'') \geq IV(w') \ \& \ p(w'') = 0))$

In a simple *should p* statement, in which there is no past operator, the temporal parameter could be provided by a (default?) speech time pronoun s^* ($\llbracket s^* \rrbracket = s^*$):

- (10) a. $\llbracket s^* [\text{should } p] \rrbracket$
 b. $\lambda w \exists w' (WA w', w, s^* \ \& \ p(w') = 1$
 $\ \& \ \neg \exists w'' (WA w'', w, s^* \ \& \ IV(w'') \geq IV(w') \ \& \ p(w'') = 0))$

Given the proposal in (7-8), *should* contrasts with *should have* in that the latter includes a past tense operator that takes scope over the modal. In the case of *should*, the speech time serves as the temporal parameter for the accessibility relation. In the case of *should have*, *have* shifts the temporal parameter to the past.

This characterization of *have* as a past temporal operator seems to make sense given the contrast between examples like (11a) and (11b):

- (11) a. She forgot to return her library book on time. But she should have returned it on time.
 b. She forgot to return her library book on time. #But she should return it on time.

We can explain the oddness of (11b) by saying that, in the absence of *have*, the temporal parameter for the accessibility relation is the speech time. At the speech time, there aren't physically accessible worlds in which she returns the book on time, the domain of quantification is empty, and the claim is infelicitous. In (11a), on the other hand, *have* shifts the temporal parameter of the accessibility relation to a past time, and there is a point in the past such that in the best worlds physically accessible at that time, she returns the library book on time (such a time will be found, presumably, before the book's due date).

The proposal in (7-8) will also be successful with more complex examples. Suppose library regulations require that she return the book next Monday, but she already did it yesterday. The time at which she was supposed to return the book lies in the future, but her actions in the past have already made such worlds inaccessible. The proposal in (7-8) correctly predicts that a *should have* form will be chosen (even though the time at which she should have fulfilled her obligation lies in the future):

- (12) a. She returned her library book yesterday. #She should return it next Monday.
 b. She returned her library book yesterday. She should have returned it next Monday.

The proposal in (7-8) predicts that (12a) is odd because there isn't a world physically accessible from the speech time in which she returns the library book next Monday. In the past, however, there were such worlds, and indeed they were the best ones. Thus (12b) is correctly predicted to be true.

4.2 *Against a Purely Temporal Perspective*

In this section we will examine some arguments against the idea that the difference between *should* and *should have* is purely temporal, as proposed by (7-8).

4.2.1 *'Have' Is Not PAST*

I will present two examples to illustrate the difficulties faced by a purely temporal analysis of the difference between *should* and *should have*. Consider a situation in which Jack, one of three candidates, has been elected as mayor after tough

elections. He was elected by a large majority, but some small procedural rules regarding the timing of the elections were ignored:

- (13) Jack should have been elected three days after the nominations were announced.

We could judge (13) true even if there isn't any time in the past in which it was true that that Jack should be elected three days after the nominations are announced. It might well be that Jack is not a better candidate than Jim or John, and the worlds in which he is elected are not better than the worlds in which one of the other candidates is elected. When we judge (13) true, we take for granted that Jack was elected, and worry only about what would have been best regarding the date. We won't obtain this result by treating *should have* as a past *should*.

Another example making the same point. There is a military parade in front of the governor's house, and it is being shown on TV. A coin is tossed and one of the soldiers is randomly chosen to be shown in a close-up on TV. Looking at him, someone could utter (14):

- (14) He should have shaved.

Suppose we judge (14) true. This will be the case even if there is no point in the past at which it is the case that he should shave. For imagine we travel backwards in time till early morning, when the soldier was deciding whether to shave or not. It is not the case that in the best worlds accessible at that time he shaves. He was randomly chosen to be on TV. Worlds in which he does not shave and is not chosen will be just as good as worlds in which he shaves and is chosen. As in the previous example, in our judgment of (14) we take for granted something that happens later (in this case, the outcome of the tossing of the coin), and this is not predicted by an analysis of *should have* as past *should*.

The examples above illustrate that the role of *have* in *should have* is not to shift the temporal parameter of the accessibility relation to the past. The examples show that the worlds that end up in the quantificational domain of *should have* are like the evaluation world at times that follow the presumed past temporal parameter for the accessibility relation. This is not predicted by a purely temporal view of the contrast.

4.2.2 'Should' and the Context Set

In this section I will show that it makes sense to link *should*-claims to our presuppositions at the speech time (the context set). Simple *shoulds* are understood as making a claim about the best worlds in the context set, whereas *should haves* may go outside the context set.

This view of the contrast between simple *should* and *should have* has important antecedents in the literature. In discussing the difference between indicative and subjunctive conditionals, Stalnaker (1975) suggested that the role of subjunctive morphology in English was to indicate that presuppositions are

being suspended, and that quantification is taking place over worlds that may be outside the context set. I will follow this intuition in the case of deontic conditionals too, and claim that the difference between *should* and *should have* is best understood by including a reference to the context set.

There is clearly a temporal component to the contrast between simple *should* and *should have*. But the contrast cannot be reduced to a temporal matter. When a speaker chooses a simple *should*, two types of information are brought together: (a) that the temporal location of the embedded clause eventuality is non-past, and (b) that we are choosing amongst worlds in the context set. The proposal I make here thus follows the spirit of Stalnaker's observation by arguing that in some cases we choose one form over the other to indicate either that we are remaining in or moving out of the context set.

Examples that show that the choice between *should* and *should have* is not purely temporal can be constructed following the pattern in (15). Imagine that a doctor is discussing the treatment of a patient under her care:

- (15) Doctor: The patient should take medicine A next Monday.
 Journalist: Mmm, actually, a mad scientist will accidentally blow up the world over the weekend.
 Doctor: Oh, ... Well,...mmm the patient should have taken medicine A next Monday.

The doctor switches from a simple *should* to *should have* as an acknowledgment of the information provided by the journalist. In switching, the doctor acts as if she accepts what the journalist says as true (though it is unclear that the journalist would actually *know* that!). Notice that the temporal semantics in (7-8) does not predict that a switch should be necessary. In the best worlds accessible from the actual world at the speech time, the scientist does not accidentally blow up the world and the patient takes medicine on Monday (I am assuming a non-deterministic view according to which it is not necessary that the scientist will blow up the world, even if it is true).

5. A Fact-Sensitive *Should*

In this section we will investigate a semantics for *should* that validates factual detachment. Our working hypothesis is that we would like a semantics for *should* statements and conditionals that tracks what happens in the actual world more closely than a Lewis-style dyadic-operator approach, so that the truth of the antecedent in a *should* conditional does lead to detachment (allowing the semantics to respond to our intuitions in (2)). At the same time, we'll need to re-evaluate the meaning of *should*, since we do not want to predict that detachment will lead to contradictions. The key, we will claim, lies in a framework that allows us to target the semantics in a more fine-grained manner by giving us access to *parts* of what is going on. The proposal will be set in the framework of situations semantics.

5.1 Preliminaries

We will talk about parts of worlds making use of the framework of situation semantics developed in Kratzer (1989, 2002, 2006). Situations are conceived as parts of worlds ('maximal situations'), and, given a Lewis-style perspective, situations are considered to be at most part of one world (the mereological 'part-of' relation will be indicated with $<$). Situations will be identified 'across worlds' via Lewis-style counterparts.

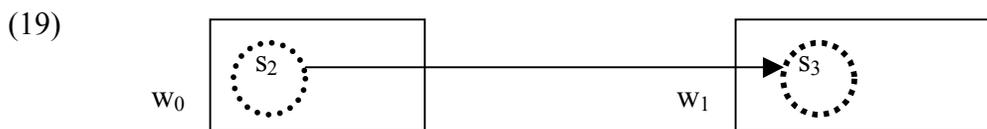
In setting up the semantics of *should*, we will augment the set of definitions with some auxiliary notions so that we can talk more easily about situations across worlds, appealing to counterparts. Given the context-dependence of the notion of *counterpart*, the auxiliary notions of *modal part of* and *modal compatibility* defined below will both be context dependent:

- (16) *modal part of*: a situation s is a modal part of ($<_m$) a situation s' iff there exists a situation s'' such that s'' is a counterpart of s and $s'' < s'$
- (17) *modal compatibility*: a situation s is compatible with a proposition α iff there exists a situation s' such that $s <_m s'$ and α is true in s' .

To build a proposal for *should* that is sensitive enough to what is going on, we will work with the idea that *should* quantifies over the best options given some facts. To track the best options associated with facts, we will make use of an auxiliary definition for a *best extension* (we will simplify things here and make the equivalent of the 'limit assumption' to talk about the best extensions of a situation):

- (18) *best-extension*: a world w is a best-extension of a situation s iff $s <_m w$ and there isn't another world w' such that $s <_m w'$ and w' is better than w (for whatever measure of goodness).

We will understand the 'better than' relation in terms of Feldman's proposal (Feldman 1986) to use a function IV to measure the intrinsic value of a possible world. Given two worlds w and w' , w' is better than w iff $IV(w') > IV(w)$. To see how (18) works, consider the schema below:



Suppose that s_2 is a situation in the actual world (w_0). The world w_1 will be a *best extension* of s_2 iff it includes a situation (let's say s_3) that is a counterpart of s_2 , and there isn't any other world that includes a situation that is a counterpart of s_2 that has a higher intrinsic value than w_1 . Notice that the definition in (18) allows for ties in goodness, and a single situation may have several *best-extensions*.

5.2 'Should' and Aspectual Heads

In Section 4 we examined (and questioned) the possibility that the difference between simple *should* and *should have* was purely temporal. Our aim in this section is to account for the Stalnaker-style generalization regarding the link between morphology and the context set presented in Section 4.2.2.

Arregui (2007) developed a proposal according to which the interpretation of aspectual heads has consequences for the interpretation of modality. The proposal had the objective of deriving the relation between aspectual morphology and the context set from the interaction between aspectual heads and events. Perfective aspect was characterized as referential aspect. It introduced a free event variable that referred to Lewis-style events and carried presuppositions with respect to the evaluation world. The link with the context set was established via diagonalization à la Stalnaker. In the proposal presented here, I will ignore the internal mechanisms that derived the link with the context set, and just assume that perfective aspect carries a presupposition with respect to the context set.

Our discussion here will be limited to the case of simple *shoulds* that embed eventive verbs and *shoulds* combined with perfect *have*. In the case of simple *shoulds* with eventive verbs, I will assume that there is a silent perfective aspectual head above a VP denoting properties of events (see Arregui 2007). I will adopt the proposal for aspectual heads found in Kratzer (1998), and modify the denotation of silent perfective aspect to include a presupposition regarding the context set in a situations framework (20):

$$(20) \quad \llbracket \emptyset_{\text{perfective}} \rrbracket^c (P) = \lambda t. \lambda s. \exists e. [P(e)(s) = 1 \ \& \ \tau(e) \subset t]$$

with the presupposition that $s < w \in c$ (the output function is defined only for situations that are part of worlds in the context set)

According to the proposal in (20), perfective aspect is responsible for situating an event in a situation at a time. The aspectual head claims that the time of the event is included within the reference time and carries the presupposition that the event is found in a world in the context set.

According to Arregui (2007), perfect aspect (contrary to perfective) does not encode information regarding the context set. Following Kratzer (1998), we will adopt a 'result state' view of the perfect (see also Parsons 1994). However, we will modify the proposal to adapt it to a situations framework:

$$(21) \quad \llbracket \text{have}_{\text{perfect}} \rrbracket (P) = \lambda t. \lambda s. \exists e. [P(e)(s) = 1 \ \& \ \tau(e) <_{\text{precedes}} t \ \& \ t \subset \tau(s)]$$

According to (21), perfect combines with a property of events and the result is a property of times true of a time t and a situation s iff the temporal location of the situation s extends (at least) until the time t and there is an event of the appropriate type in the situation that precedes the time t .

I will follow Arregui (2007) (and others) in allowing the modal to set the temporal parameter of the embedded clause. *Should* will locate the temporal argument of the embedded clause at a non-past time. When *should* combines with

simple eventive predicates, the time of the event will be non-past (future). When *should* combines with perfects, the result state will include a non-past time (allowing the event that gave rise to the result state to be past, present or future). For example, in the case of *He should take medicine A on Monday*, the proposition manipulated by the modal will be (22) (where t is a non-past time):

(22) $\lambda s: s <_w \in c. \exists e. [\text{He-takes-medicine-A-on-Monday}(e)(s) = 1 \ \& \ \tau(e) \subset t]$

The view of aspect adopted here will have consequences for the interpretation of deontic sentences. There will be a difference between simple *should* claims that embed eventive verbs, and *shoulds* that embed perfects. With eventive clauses, the embedded proposition can only be true in worlds in the context set. The presuppositions of perfective aspect will place restrictions on the cases quantified over. If we know that the embedded clause is false, then the embedded proposition will not be true in any world in the context set. The prediction is that the corresponding *should* statement will be false (though it might be more accurate to say it would be infelicitous). In the case of perfects there is no link between the embedded proposition and the context set. The truth of a *should have* statement will depend only on the constraints imposed by the modal. To see an example of this, consider again the doctor who has just found out that the world will end over the weekend. In a context in which this is presupposed, there will be no worlds in the context set in which the patient takes medicine A on Monday. Given the semantics for perfect aspect, it would be infelicitous to say *He should take medicine A on Monday*. It will be necessary to switch to a perfect (unrelated to the context set) and claim that *He should have taken medicine A on Monday*.

5.2 On the Interpretation of 'Should'

To understand the meaning of *should* we need to understand how the domain of quantification of the modal is identified. It is clear that in the case of deontic modality we can end up quantifying over worlds that we know are different from the actual world (the clause embedded under the modal can be simply false in the actual world). But it is also clear that deontic claims are (can be) contingent, and that the set of worlds quantified over depends on what is happening in the actual world. Deontic claims are made true by things happening in the actual world.

The relation between actual world facts and the truth of deontic claims can appear particularly intuitive in the case of reparational duties. Consider (23):

(23) She returned the book late. She should have paid a fine.

Our judgments regarding the truth of the deontic claim in (23) are closely tied to the fact that she returned the book late. In a sense, it is this feature of the actual world that 'justifies' the modal claim. But we cannot say that the late return of the book is all that matters. Judgments regarding deontic statements can be sensitive to a variety of things. Suppose she is a careless, wealthy supermodel, who

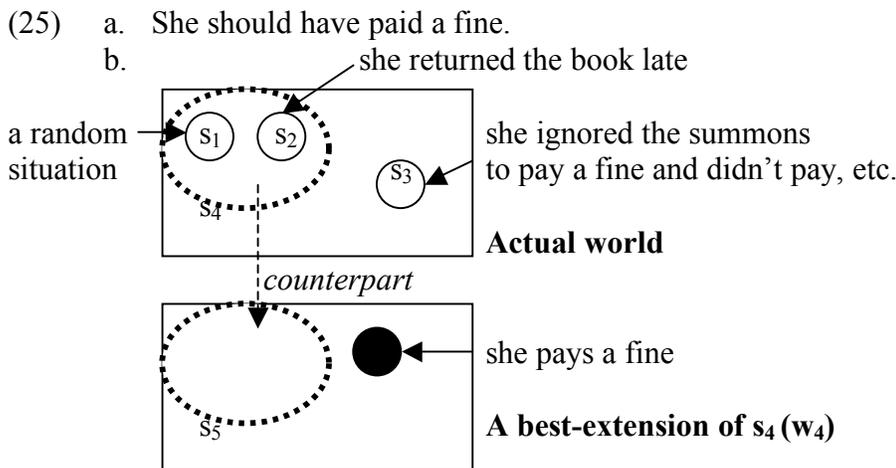
constantly returns books late and simply does not care. Ok, she should pay a fine. But suppose she is a poor, hardworking grad student, who slipped up once, is very sorry, and would be completely ruined by a fine and have to leave school. Our intuitions could waver. Taking into account only library regulations we would say, go ahead, fine her. But our judgements are often more complex, the measure of goodness can be an 'overall' measure, sensitive to various factors. Library regulations can be overridden by more important considerations. In judging (23), we do not only pay attention to the fact that she returned the book late, we also worry about the interaction between that fact and the other things going on in the world.

The proposal for *should* that is presented in (24) below addresses these concerns by tracking the existence of situations in the world that 'justify' modal claims and worrying about their interaction with other features:

- (24) $\llbracket \text{should } \alpha \rrbracket$ is true in w iff
 $\forall s < w$ that are α -compatible,
 $\exists s' [s < s' \ \& \ \forall w' \text{ that are best-extensions of } s', \alpha(w') = 1]$

According to (24), *should* α will be true iff all the situations compatible with α are such that there is an extension in the actual world such that in its best extensions, α is true. Quantification takes place over the situations in the actual world that are compatible with the embedded clause. By appealing to situations, we can identify the features in the world that 'stay the same' in the worlds quantified over and at the same time we can ignore the features of the world incompatible with the modal claim. Quantification takes place over all the situations in the actual world compatible with the embedded clause, thus ensuring that all the relevant facts in the world can be weighed in. And we worry about the existence of facts in the world that 'justify' the modal claim (thus tracking the situations that make the modal claim true):

We will investigate how (24) works by discussing an example. Suppose she returned the book late, and ignored the notice to pay a fine.



The sentence in (25a) will be true in the actual world iff for every situation in the actual world that is compatible with her paying a fine (that is, for all situations excluding s_3) there is an extension such that in its best extensions, she pays a fine. Let us consider the case of s_1 as an example. In the actual world, there is an extension of s_1 (for example, s_4), such that in all its best extensions, she pays a fine. Remember that a world will be a best extension of s_4 iff it includes a counterpart of s_4 and is amongst the best worlds to do so. So, w_4 will be a best extension of s_4 if, for example, s_5 is a counterpart of s_4 and there isn't any other world that also includes a counterpart of s_4 that has a higher intrinsic value than w_4 .

By quantifying over all (compatible) situations and appealing to extensions, the proposal in (24) captures the factual dependency of modal claims. If a modal claim is justified by something that has happened in the world (let us say, situation s), then for all situations s' in the world there will be an extension (s' with s) that will have best extensions in which the modal claim is true (for reasons of space, I will not be able to discuss different kinds of examples, the reader is referred to Arregui (2008) for a more thorough presentation).

The 'overall' nature of the measure of goodness is captured in (24) by the universal quantification over parts of the world. Let us consider again example (25b), and the case in which she was rich and careless vs. the case in which she was poor but honest. If she returned the book late, then for all situations in the actual world there will be an extension in which she returned the book late. If she is rich and careless (and the measure of goodness is sensitive to the satisfaction of library regulations and social justice) then all situations compatible with her paying a fine will have extensions that will be improved if she does pay a fine. And so in the best extensions, she will pay a fine. On the other hand, if she is poor and honest, then for some situations in the actual world (corresponding to her vulnerable financial state and honest nature) it won't be the case that there is an extension such that in all its best extensions she pays a fine. It will be true that for all situations there is an extension in which she returns the book late. But it won't be true that in all the best-extensions of this situation she pays a fine. It may be that in the best extensions of the actual world situation corresponding to her poverty, honesty, and late book she is not fined, and gets away with a warning. These situations won't be ideal from the point of view of the (library) law, but the outcome will reflect the variety of criteria that come together when we evaluate relative goodness.

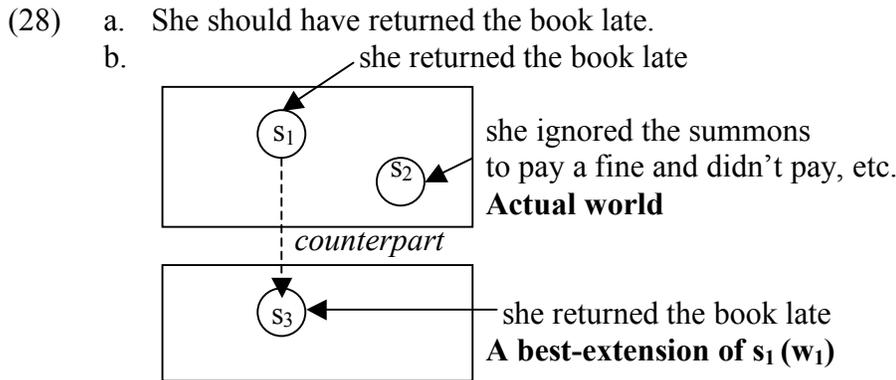
Given the existential quantifier in the denotation of *should*, the proposal in (24) predicts that if she borrowed a book, returned it late, and refused to pay a fine, both statements in (26) can be true:

- (26) a. She should have returned the book on time.
b. She should have paid a fine.

This is because in the actual world, there is a situation that justifies (26a) (that she borrowed the book) and there is a situation that justifies (26b) (that she returned the book late). The predictions made by (24) in this case do not necessarily vary

from the predictions that we would obtain with a temporally sensitive accessibility relation as we saw in Section 4.1. According to the proposal in (7-8), *have* shifts the temporal parameter of the accessibility relation to some point in the past. Given the history of the world, there was a point in the past such that in the best worlds accessible from there she returned the book on time (a point after she borrowed the book and before it was due). So (26a) will be true. And there is also a point in the past such that in the best worlds accessible from there she paid a fine (a point after she returned the book late and before she refused to pay the fine). So (26b) will be true.

As a last remark in this section, let me note a polemic side effect to the proposal presented here. If we have a statement of the form *should* α , and α is true, *should* α will be trivially true. If there are features in the world that make α true, then it will be trivial to find in the world a situation such that in all of its best extensions, α is true, and this will make *should* α true. In a sense, the truth of α collapses the semantic machinery built around *should*. This is illustrated below:



The semantics we have proposed for *should* is closely tied to the facts. It looks for the best alternatives given the facts. For any particular thing that has happened, its best alternatives will include it. Thus, if it is true that she returned the book late, then for all the situations in the world, there will be an extension (to a situation that includes the late return of the book), such that in its best extensions, she returned the book late. For that situation, it will be true in all of its extensions that she returned the book late, therefore, it will also be true in its best extensions.

The fact that the truth of the complement makes the truth of the deontic statement trivial is something that is problematic, and I will not be able to resolve here. I will propose the stop-gap solution of excluding from the domain of quantification of *should* situations that themselves make the embedded statement true. A revised semantics for *should* would then be as follows:

- (29) $\llbracket \text{should } \alpha \rrbracket$ is true in w iff
 $\forall s < w$ that are α -compatible and such that α is not true in s ,
 $[\exists s' [s < s' \ \& \ \forall w' \text{ that are best extensions of } s', \alpha(w') = 1]]$

With this addition, the accidental truth of α will not have consequences regarding whether *should* α is true. The situations that make α true will be ignored, and the

truth value of the statement will truly depend on whether α is the better option or not.

5.3 Conditioning 'Should'

I will present a proposal for conditioning *should* by building on an example. Consider the conditional in (29):

- (29) If he had forgotten to pay his taxes (last May), he should have been fined \$200.

We'll worry only about the value of possibilities as defined by the law. Suppose that according to the law there is a link between the fine one receives for being late with the taxes and one's salary. In his case, he should have been fined \$200 because he earns \$20 000 per year. Suppose also that as a matter of fact he did pay his taxes and was not fined \$200.

The truth of (29) seems closely tied to the fact that he earns \$20 000 per year (this is an actual world fact, not an ideal). We will build the interpretation of the conditional on the intuition that, if we take this fact, and then add that he forgot to pay his taxes, in the best circumstances (given value as characterized by the law!) he is fined \$200. The analysis we are looking for is one where the *if*-clause restricts *should*, but the modal is still able to access what is actually going on. A proposal with this objective is given in (30):

- (30) $\llbracket \text{if } p, \text{ should } q \rrbracket$ is true in w iff
 $\forall s <_w$ that are q -compatible (and in which q is not true)
 $[\exists s' [s < s' \ \& \ \{s'' : s'' \text{ is a } p\text{-minimal extension of } s'\} \subseteq \{s'' : \forall w' : w' \text{ is a best extension of } s'', q \text{ is true in } s''\}]$

The proposal in (30) treats *if*-clauses as restrictors, and at the same time allows specific facts to have an effect on the semantics of the conditional. According to (30), a conditional will be true iff for all the (relevant) situations in the world, there is an extension in the world such that all the minimal extensions that make the antecedent true have best extensions in which the consequent is true.

The proposal in (30) makes use of an auxiliary notion (*p-minimal extension*) defined in (31):

- (31) *p-minimal extension*
 Given a situation s and a proposition p , s' is a p -minimal extension of s iff $s <_m s'$ and p is true in s' and there isn't a situation s'' such that $s <_m s'' <_m s'$ and p is true in s'' and $s'' \neq s'$.

The claim made in (30) is that when evaluating a deontic conditional of the form *if p, should q*, we worry about the minimal extensions of actual world situations that make the antecedent true. We focus on minimal extensions to exclude from

the domain of quantification situations in which the consequent is false. To see why this is important, consider again the case above. We have found an actual world situation that 'justifies' the fine being \$200 (he earns \$20 000). To evaluate the conditional, we extend it to a situation in which he forgot to pay his taxes (a situation in which the antecedent is true). If we worry about all extensions in which he forgot to pay his taxes, we will include in the set of antecedent situations, situations in which he was not fined \$200 (the tax office forgot him, or they charged him more than they should have, or...). Those situations won't have best-extensions in which he was fined \$200! If we allow this kind of situations in, the machinery put in motion to see what is better will be disarmed.

6. Conclusion

To conclude, let us review some of the results we have achieved. We have set up the semantics of *should* in such a way that we predict that if she forgot to return her library book on time, both the claim that *she should have returned the book on time* and the claim that *she should have paid a fine* can come out true (this prediction is also made by a view of *have* as a past operator taking scope over *should*, as we saw in Section 4.1). Given the epistemic distinction we have established between simple *should* and *should have* (arising from the semantics of perfective and perfect aspect), if we know that she returned the book two days ago, when she should have returned it tomorrow, the claim that *she should return the book tomorrow* will not be felicitous (there is no world in the context set in which she does return it tomorrow).

To see the predictions made by our proposal with respect to detachment, let us consider a (simplified) library book version of Chisholm's sentences:

- (32)
- a. She should return the library book on time.
 - b. If she returned the library book late, she should pay a fine.
 - c. She returned the library book late.
 - d. She should have returned the library book on time.

We will only be able to utter (32a) felicitously in a context in which we do not know that she returned the library book late (maybe we were unaware of the book's due date). This is predicted because if we know that she returned the book late, then there won't be any worlds in the context set in which she returns the library book on time (the presuppositions of the silent perfective aspectual head won't be satisfied). On the other hand, knowing that she returned the library book late, we will be able to felicitously utter (32d), and it could well be true (it would depend on what else had been happening in the world and on the measure of value that was contextually relevant). The truth of (32b) and (32c) will lead us to the conclusion that she should pay a fine. Given the proposal in (20), (32b) will be true only if both the propositions that she return the book late and pay a fine are compatible with what we know, and if all the situations in the world compatible with her paying a fine are such that they have an extension in the world, such that

for the minimal extensions in which she returns the book late, the best extensions are such that she pays a fine. If she borrowed the book, then all (relevant) situations in the world will have extensions to situations in which she borrowed the book, and in the minimal extensions of those in which she returned the book late, it will be true that the best extensions are such that she pays a fine. If she actually returned the book late (32c), then some of those minimal antecedent extensions will be in the actual world. And since all those minimal antecedent extension situations have best extensions in which she pays a fine, then there will be a situation in the actual world (fitting all the conditions) such that in all its best extensions, she pays a fine. So, she should pay a fine.

References

- Åqvist, Lennart: 2002, 'Deontic Logic', in D.M. Gabbay and F. Guentner (eds.) *Handbook of Philosophical Logic*, vol. 8, Kluwer, Dordrecht, 147-264.
- Arregui, Ana: 2007, 'When Aspect Matters: The Case of *Would*-Conditionals', *Natural Language Semantics* **15**, 221-264.
- Arregui, Ana: 2008, 'Detaching conditions from obligations', ms., U. of Ottawa.
- Carmo, José and Andrew Jones: 2002, 'Deontic Logic and Contrary-to-Duties', in D.M. Gabbay and F. Guentner (eds.) *Handbook of Philosophical Logic*, 2nd edition, vol. 8, Kluwer, Dordrecht, 265-343.
- Chisholm, Roderick: 1963, 'Contrary to Duty Imperatives and Deontic Logic', *Analysis* **24**, 33-36.
- Copley, Bridget: 2006, 'What should *should* mean', ms., CNRS/ Paris 8.
- Feldman, Fred: 1986, *Doing the Best We Can: An Essay in Informal Deontic Logic*, Reidel, Dordrecht.
- von Fintel, Kai and Sabine Iatridou: 2006, 'How to Say *Ought* in Foreign', MIT.
- Ippolito, Michela: 2003. 'Presuppositions and Implicatures in Counterfactuals', *Natural Language Semantics* **11**, 145-186.
- Frank, Anette: 1996, *Context Dependence in Modal Constructions*, Doctoral Dissertation, Universität Stuttgart.
- Kratzer, Angelika: 1989, 'An Investigations of the Lumps of Thought', *Linguistics and Philosophy* **12**, 607-653.
- Kratzer, Angelika: 1998, 'More Structural Analogies Between Pronouns and Tenses', in D. Strolovitch and A. Lawson (eds.). *Proceedings of Salt VIII*, CLC Publications, Ithaca.
- Kratzer, Angelika: 2002, 'Facts: Particulars of Information Units?', *Linguistics and Philosophy* **25**, 655-670.
- Kratzer, Angelika: 2006, 'Situations in Natural Language Semantics', in *Stanford Encyclopedia of Philosophy*, (http://plato.stanford.edu/entries/situations_semantics/#SitSemImpDomRes).
- Lewis, David: 1973. *Counterfactuals*, Harvard University Press, Cambridge.
- Parsons, Terence: 1994, *Events in the Semantics of English: A Study in Subatomic Semantics*, MIT Press, Cambridge.
- Stalnaker, Robert: 1975, 'Indicative Conditionals', *Philosophia* **5**, 269-286.