On the Agent-Relativity of ‘Ought’

JUNHYO LEE

Forthcoming in Analysis

Abstract: In the standard theory of deontic modals, ‘ought’ is understood as expressing a propositional operator. However, this view has been called into question by Almotahari and Rabern’s puzzle about deontic ‘ought’, according to which the ethical principle that one ought to be wronged by another person rather than wrong them is intuitively coherent but the standard theory makes it incoherent. In this paper, I take up Almotahari and Rabern’s challenge and offer a refinement of the standard theory to handle the puzzle. I propose that ‘ought’ is evaluated relative to contextual parameters (e.g., Kratzer’s conversational backgrounds, Finlay and Snedegar’s alternative sets) and those contextual parameters are sensitive to agents as well as possible worlds.

Keywords deontic modality; agents; ought to do; ought to be; obligations; centered worlds

1. Almotahari and Rabern’s argument

In the standard theory of deontic modals (Chisholm 1964, Williams 1981, Kratzer 1991, 2012, von Fintel 2012, Lee 2021), ‘ought’ is regarded as expressing a propositional operator. On this view, a sentence of the form ‘Ought φ’ is true iff φ is true in all suitably restricted ideal worlds. The relevant ideal worlds are determined by two premise sets (= sets of propositions1) (i.e., what Kratzer calls conversational backgrounds): the modal base f and the ordering source g. The modal base f determines which worlds are accessible from the base world w. This preliminary domain is represented as ∩f(w) (i.e., the set of worlds in which all propositions in f(w) are true). The ordering source g induces an ordering on the worlds in ∩f(w). Thus, ‘Ought φ’ can be glossed as ‘All the

1 Propositions are understood as sets of possible worlds.
worlds in \( \cap f(w) \) that are best ranked by \( g(w) \) are \( \phi \)-worlds’. This analysis is simple but has been proved to be empirically fruitful while preserving the uniformity of ‘ought’.

However, Almotahari and Rabern (2023) have recently presented an interesting puzzle about deontic ‘ought’, which raises a significant challenge against the standard theory of ‘ought’. The puzzle is that the following ethical principle from ancient Greek philosophy sounds intuitively coherent, but the standard theory of ‘ought’ makes it incoherent.

(1) One ought to be wronged by another person rather than wrong them.\(^2\)

Assuming that ‘rather than’ denotes a binary relation >, the meaning of (1) can be analysed as (2).

(2) \( \forall x \forall y [x \neq y \supset \text{Ought} <f(w),g(w)> (Wyx > Wxy)] \)\(^3\)

And given that \( a \) and \( b \) are distinct agents, (2) entails (3) and (4).

(3) \text{Ought}<f(w),g(w)> (Wba > Wab)

(4) \text{Ought}<f(w),g(w)> (Wab > Wba)

According to the standard theory of ‘ought’, ‘Ought<\( f(w),g(w) > \phi > \psi \)’ is true iff, among the \( \phi \)-or-\( \psi \)-worlds in the relevant domain \( \cap f(w) \), the highest ranked worlds relative to \( g(w) \) are \( \phi \)-worlds and not \( \psi \)-worlds (or equivalently, the highest ranked \( \phi \)-worlds are more highly ranked

---

\(^2\) Almotahari and Rabern (2023) credit Müller (1977) for the example, and Müller (1977) attributes the example to Democritus.

\(^3\) In the standard theory, ‘ought’ is evaluated relative to conversational backgrounds such as the modal base \( f \) and the ordering source \( g \). But there are two ways to formalize this idea. Conversational backgrounds can be represented as parameters at the index in Lewisian terms (Lewis 1980) or the circumstance of evaluation in Kaplanian terms (Kaplan 1989). (1) illustrates the representation of conversational backgrounds as parameters at the level of meta-language. On the other hand, they can be represented as arguments at the level of object language. (2) is an example in which conversational backgrounds are represented as arguments of ‘ought’. (1) \( \text{⟦Ought} \phi \text{⟧}_{f,g,w}^{f,g,w} \) is true iff \( \forall w' \in \text{BEST}(f, g, w). \text{⟦} \phi \text{⟧}_{f,g,w}^{f,g,w} = 1. \) (Kratzer 2012: Ch2, Portner 2009)

(2) a. \( \text{⟦Ought} = \lambda f. \lambda g. \lambda \phi. \lambda w. \forall w' \in \text{BEST}(f, g, w). \text{⟦} \phi \text{⟧}(w') = 1. \) (Kratzer 2012: Ch1, von Fintel & Heim 2020)

b. \( \text{⟦Ought}<f(w),g(w)> \phi \text{⟧} \) is true iff \( \forall w' \in \text{BEST}(f, g, w). \text{⟦} \phi \text{⟧}(w') = 1. \)

In this paper, I adopt the second approach, because it is better suited to explaining some troublesome data presented by Schroeder (2011). Thus, if one sympathizes with Schroeder’s data, they will have a good reason to favor the second approach over the first. I will discuss Schroeder’s data in \( \S 3 \).
than the highest ranked ψ-worlds)⁴ (Almotahari & Rabern 2023). If ‘Ought (Wba > Wab)’ is true, among the Wab-or-Wba-worlds in ∩f(w), the highest ranked worlds relative to g(w) are Wba-worlds and not Wab-worlds. Also, if ‘Ought (Wab > Wba)’ is true, among the very same worlds, the highest ranked worlds relative to g(w) are Wab-worlds and not Wba-worlds. Thus, if (3) and (4) are both true, the highest ranked worlds relative to g(w) are Wba-worlds and not Wab-worlds. But this is a contradiction. Therefore, (3) and (4) are inconsistent.

The problem with the standard theory is that while ‘ought’, as a propositional operator, appears to be suitable for expressing how things ought to be, it fails to capture what actions someone ought to do. In the standard theory, ‘ought’ does not take into account the source of a wronging. It only cares about how the worlds are ranked. So, it does not matter whether it’s a wronging b or b wronging a.

To put the point differently, the standard theory falls short in that it fails to distinguish (5) from (6), since in the standard theory they are all analysed as (3). That is, the standard theory is not fine-grained enough to capture the difference between (5) and (6).

(5) Ann ought to be wronged by Betty rather than wrong her.

(6) a. It ought to be that Ann is wronged by Betty rather than wrongs her.

   b. It ought to be that Betty wrongs Ann rather than Ann wrongs Betty.

The most natural interpretation of (5) is agentive;⁵ it describes an action that Ann, as an agent, ought to perform. So, on this reading, (5) can be paraphrased as ‘Ann is obliged to, or has

---

⁴ Given the limit assumption, this analysis is equivalent to Kratzer’s analysis of ‘φ is as good a possibility as ψ’ (Kratzer 1991: 644).

⁵ Although there is no consensus on exactly how to distinguish between agentive and non-agentive ‘ought’s, Schroeder’s (2011) five hallmarks of agentive ‘ought’ (deliberative ‘ought’ in his terms) can provide a rough guideline on how to differentiate between them: (i) deliberative ‘ought’s matter directly for advice. When someone comes to you with the question of what to do, if you know what s/he deliberatively ought to do, you know what is advisable for him/her. (ii) Deliberative ‘ought’s settle practical deliberation. If you know what one deliberatively ought to do, then that settles the question of what to do. (iii) Deliberative
most reason to, be wronged by Betty rather than wrong her’. On the other hand, the most natural interpretation of (6) is non-agentive;' it describes a state of affairs that ought to be the case. On this reading, (6) can be paraphrased as ‘It is best or most desirable that Ann is wronged by Betty rather than wrongs her’.7 Intuitively, (5) is a logical consequence of (1) but (6) is not. The basic idea behind (1) is that one ought to endure a wronging rather than be its agent, because wronging someone is ultimately damaging to one’s own character or well-being. From this, it follows that Ann ought to endure a wronging rather than be its agent. However, even if (1) is true, (6) is not true. This is because (6) expresses what is good from the perspective of the world. Given that either Ann wrongs Betty or Betty wrongs Ann, it does not matter from the perspective of the world whether the agent of wronging is either Ann or Betty.

2. Agent-relative conversational backgrounds

In this section, I argue that the standard theory, once refined, can accommodate the coherence of (1) while maintaining the uniformity of ‘ought’. For concreteness, I will use Kratzer’s notion of conversational background to present my solution, but it should be noted that it can be adapted to any theory with a device similar to Kratzer’s conversational backgrounds such as alternative sets.

‘ought’s are closely related to the notion of accountability. If an agent S deliberatively ought to do some action X, S is accountable if S fails to perform X. (iv) The deliberative ‘ought’ implies the ability ‘can’. If an agent S deliberatively ought to do some action X, it has to actually be in S’s power to do X. (v) Deliberative ‘ought’s are closely related to the notion of obligation.

6 One might claim that (6) does have an agentive interpretation (Finlay & Snedegar 2014). But it seems obvious that the most natural interpretation of (6) or the interpretation of (6) in regular contexts is non-agentive, so in this paper I will focus on their non-agentive interpretation and remain neutral on whether (6) can semantically express an agentive reading. For more information, see Schroeder 2011 and Finlay & Snedegar 2014.

7 Note that the distinction between agentive and non-agentive ‘ought’s roughly corresponds to Schroeder’s (2011) deliberative and evaluative ‘ought’s, Humberstone’s (1971) agent-implicating and situational ‘ought’s, Feldman’s (1986) ‘ought to do’s and ‘ought to be’s, Price’s (2008) agential and situational ‘ought’s, Finlay & Snedegar’s (2014) agential and non-agential ‘ought’s, and Chrisman’s (2016) and Skibra’s (2022) agentive and non-agentive ‘ought’s.
in contrastive semantics (Snedegar 2012, Finlay 2014, Finlay & Snedegar 2014, Cariani 2016, Lassiter 2017). In the standard theory, ‘ought’ is evaluated relative to conversational backgrounds, and conversational backgrounds are sensitive to possible worlds. I propose that conversational backgrounds are sensitive to agents as well as worlds.8,9

The basic idea is that ‘ought’ can be associated with two types of conversational backgrounds: world-relative and agent-relative conversational backgrounds. While world-relative conversational backgrounds are functions of worlds and give rise to non-agentive interpretations, agent-relative conversational backgrounds are functions of agents and give rise to agentive interpretations. It is well known that ‘ought’-statements can have a wide range of modal flavors, including epistemic, stereotypical, deontic, bouletic, teleological, lawful, prudential, etc. The standard theorists capture this feature by arguing that a variety of modal flavors are determined by the modal base f and the ordering source g. For example, Kratzer says that a modal has a deontic reading when it takes a circumstantial modal base (= a set of relevant facts) and a deontic ordering source (= a set of moral principles). However, Kratzer does not distinguish between agentive and

---

8 The main motivation for world-relativity of conversational backgrounds is that modal sentences can express contingent propositions (Kratzer 2012: 99). If so, one might worry that if conversational backgrounds are agent-relative, modal sentences are not contingent anymore. But this issue can be resolved if we accept a Lewisian ontology (Lewis 1986). In a Lewisian ontology, each individual belongs to a unique world. If we do not want to make this assumption, we can understand agent-relative conversational backgrounds as functions of centered worlds consisting of a world, an individual, and a time. For further discussion, see fn.11.

9 There is some independent reason to believe that modal bases/ordering sources can have an individual argument. Kratzer (2012: Ch1) argues that a modal has two arguments: a restriction and a scope. The restriction argument usually is implicit and provided by the context of utterance. But it may be overtly realized as an ‘in view of’-phrase, as in (1).

(1) In view of one’s obligations (= the restriction argument/conversational background), one ought to be wronged by another rather than wrong them (= the scope argument/prejacent).

(2) For every x, Ought-x: x is wronged by another rather than wrong them.

(1) can be analysed as (2), where ‘one’ is interpreted as a free variable over individuals and this free variable is bound by a higher quantifier such as the default universal quantifier (Epstein 1984) or the generic operator GEN (Moltmann 2006, 2010, Bhatt & Pancheva 2017). If this is correct, the data like (1) requires that the ordering source should at least sometimes be a function of an individual.
non-agentive deontic interpretations. To capture this distinction, I propose that circumstantial modal bases and deontic ordering sources are divided into two types: agentive and non-agentive.

In this picture, if a modal takes an agentive modal base and an agentive ordering source, it has an agentive interpretation. If it takes a non-agentive modal base and a non-agentive ordering source, it has a non-agentive interpretation. The main difference between agentive and non-agentive conversational backgrounds is that while non-agentive conversational backgrounds are functions of worlds, agentive conversational backgrounds are functions of agents. More specifically, I propose the following as the conversational backgrounds for agentive and non-agentive ‘ought’s:

**World-Relative Conversational Backgrounds**

\[ \cap f(w): \{ w' : w' \text{ is compatible with the circumstances of } w \} \]

\[ g(w): \{ p: p \text{ is what is good at } w \} \]

**Agent-Relative Conversational Backgrounds**

\[ \cap f(x): \{ w' : w' \text{ is compatible with } x's \text{ circumstances and abilities} \}^{10} \]

\[ g(x): \{ p: p \text{ is one of } x's \text{ moral obligations} \} \]

---

10 In this paper, I focus on cases in which the ordering source parameter needs to be agent-relative. However, I have observed in (2022: ch.2) that ‘can’, as well as ‘ought’, has agentive and non-agentive interpretations. On the agentive reading, ‘can’ is naturally interpreted as ‘be able to’ or ‘have the power/ability to’, and so (1) can be paraphrased as ‘Bill is able to lift this table’. On the other hand, on the non-agentive reading, ‘can’ is naturally interpreted as ‘it is possible that’, and so (2) can be paraphrased as ‘It is possible that Larry wins the lottery’.

(1) Bill can lift this table (agentive)

(2) Larry can win the lottery (non-agentive)

On Kratzer’s (1991) account, ‘can’ has a circumstantial modal base and an empty ordering source. Since Kratzer posits only one kind of circumstantial modal base, her theory is not fine-grained enough to capture the two distinct readings of ‘can’. However, if my analysis is correct, the distinction between agent-relative and world-relative modal bases can be employed to capture these intuitive readings. That is, it can be proposed that while the non-agentive ‘can’ is characterized by a world-relative circumstantial modal base and an empty ordering source, the agentive ‘can’ is characterized by an agent-relative circumstantial modal base and an empty ordering source.
As can be seen from the above, while non-agentive conversational backgrounds are functions of worlds, agentive conversational backgrounds are functions of agents. This difference leads to a difference in interpretations. A non-agentive ordering source is a set of propositions representing what is good or desirable from the perspective of the base world, so it includes propositions like *there is no wrongdoing*. On the other hand, an agentive ordering source is a set of propositions representing an agent’s obligations, so it includes propositions like *x does not wrong anyone*.

This idea can be used to explain agentive and non-agentive ‘ought’s. On my view, (5), repeated below, can be evaluated relative to either a world-relative or agent-relative conversational background.

(5) Ann ought to be wronged by Betty rather than wrong her.

Suppose that a world-relative ordering source g(w) is \{There is no wrongdoing\} and an agent-relative ordering source g(Ann) is \{Ann does not wrong anyone\}. When (5) is evaluated relative to g(w), it is not true. (5) is true relative to g(w) iff among the Ann’s wronging Betty or Betty’s wronging Ann-worlds, the worlds in which Ann is wronged by Betty and so Betty wrongs Ann are more highly ranked than the worlds in which Ann wrongs Betty. Since g(w) requires that there be no wronging at all, as long as a world involves any wronging, it does not matter whether it is Ann’s wronging Betty or Betty’s wronging Ann. That is, Ann’s wronging Betty-worlds and Betty’s wronging Ann-worlds are equally ranked. On the other hand, when (5) is evaluated relative to g(Ann), it is true. This is because the highest ranked worlds determined by g(Ann) are where Ann does not wrong anyone. On this interpretation, it *does* matter whether it is Ann’s wronging Betty or Betty’s wronging Ann. That is, as far as Ann’s obligations are concerned, the worlds where
Betty wrongs Ann and so Ann endures a wronging are more highly ranked than the worlds where Ann is the agent of wronging Betty.

Let’s now see how my proposal solves Almotahari and Rabern’s puzzle. On my view, agentive ‘ought’s take agent-relative conversational backgrounds, so the meaning of (1) should be analysed as (2’) rather than (2). And (2’) entails (3’) and (4’), which are consistent.

(1) One ought to be wronged by another person rather than wrong them.

(2’) ∀x∀y[x ≠ y ⊃ Ought<fx,gx> (Wy > Wxy)]

(3’) Ought<fa,ga> (Wba > Wab)

(4’) Ought<fb,gb> (Wab > Wba)

The truth-condition for ‘Ought<fx,gx> (φ > ψ)’ can be given as: among the φ-or-ψ-worlds in the relevant domain ∩f(x), the highest ranked worlds relative to g(x) are φ-worlds and not ψ-worlds. If (3’) is true, among the Wab-or-Wba-worlds in ∩f(x), the highest ranked worlds relative to a’s obligations are Wba-worlds and not Wab-worlds. If (4’) is true, among the very same worlds, the highest ranked worlds relative to b’s obligations are Wab-worlds and not Wba-worlds. (3’) and (4’) can both be true, because while the highest ranked worlds relative to a’s obligations are where a endures b’s wronging, the highest ranked worlds relative to b’s obligations are where b endures a’s wronging.

Also, my proposal captures the difference between (5) and (6). On my view, conversational backgrounds can take either a possible world or individual as their argument. (5), on its most natural reading, is associated with (7a), where conversational backgrounds take an individual as their argument, and so ‘ought’ has an agentive interpretation. On the other hand, (6), on its most natural reading, is associated with (7b), where conversational backgrounds take a world as their argument, and so ‘ought’ expresses a non-agentive reading.
(7) a. Ought_{\langle f(Ann),g(Ann) \rangle} (Ann is wronged by Betty rather than wrongs her)

b. Ought_{\langle f(w),g(w) \rangle} (Ann is wronged by Betty rather than wrongs her)

Before closing the section, it is important to note that the analysis developed here should be distinguished from the view that agentive ‘ought’ expresses a relation between an agent and a proposition (e.g., Wedgwood 2006, Schroeder 2011), according to which the structure of ‘ought’ can be formulated as either ‘Ought_{\langle a \rangle}(p)’ or ‘Ought (a, p)’, where ‘a’ refers to an agent. On this view, agentive ‘ought’ is like ‘believe’, ‘want’, or ‘try’ in that it directly takes an agent as one of its arguments. In contrast, my view treats agentive ‘ought’ as a propositional operator that does not directly take an agent as an argument. Instead, the contextual parameters for ‘ought’ may be functions of agents as well as worlds. These are two different ways to encode agent-relativity in the semantics of ‘ought’.

3. The problem of overgeneration

On my view, ‘ought’ can be evaluated relative to either a world-relative or agent-relative conversational background. If so, my view seems to predict that (6) can have both agentive and

---

11 I take the lesson from Almotahari and Rabern’s puzzle to be that agent-relativity needs to be encoded in the semantics of ‘ought’. The key idea of this paper is that agent-relativity can be encoded as an argument of contextual parameters for ‘ought’ (e.g., Kratzer’s conversational backgrounds), rather than a direct argument of ‘ought’. But there are at least three different ways to implement this idea: (i) Conversational backgrounds may freely take either a world or an individual. (ii) Conversational backgrounds may freely take either a world or a centered world, which is a <world <individual, time>>-pair (Lewis 1979, Egan 2006, Liao 2012). (iii) Conversational backgrounds may freely take either a boring centered-world or an interesting centered-world. It can be proposed that a boring centered world does not have a substantial individual component, so it is of the form <w,<w,t>>, representing a worldly perspective. On the other hand, an interesting centered world has a substantial individual component, so it is of the form <w,<i,t>>, representing an agential perspective. I take these three implementations as being in the same spirit. But for simplicity, I adopt the first implementation in the paper. I thank two anonymous referees for bringing this point into my attention.

12 Chrisman 2012 and Lee 2021 present some good reasons to favor a propositional operator approach over a relational approach.
non-agentive readings, depending on which conversational background ‘ought’ takes. However, Schroeder (2011) claims that while (5) has an agentive reading for Ann, (6) does not. If so, the analysis developed here appears to face the problem of overgeneration. That is, it generates readings that are not available.

I would like to point out two things. First, there is no consensus on whether or not (6) semantically conveys an agentive interpretation. For example, although I will not discuss their examples in detail, Finlay and Snedegar (2014) claim that at least in some special contexts, (6) may have agentive interpretations. If their judgments are correct, my analysis does not need to be revised.

Second, even if Schroeder’s judgments are correct, the analysis developed here, when combined with Heim and Kratzer’s (1998) theory of movement, can explain why (5) has an agentive reading for Ann but (6) does not. Let’s first consider how (5) can have an agentive interpretation. At the beginning, ‘Ann’ in (5) is base-generated as the subject of the main verb ‘be wronged by Betty rather than wrong her’. At this point, the structure of (5) looks like (8), where ‘π’ is a neutral variable over worlds or individuals.

\[
(8) \left[ TP/IP \text{ Ought}_{\langle r(\pi), g(\pi) \rangle} \left[ VP \text{ Ann [is wronged by Betty rather than wrongs her]} \right] \right]
\]

Then, ‘Ann’ moves up from the subject position of the main verb (= the specifier (Spec) of the verbal phrase (VP)) to the subject position of ‘ought’ (= the specifier (Spec) of the tensed or inflectional phrase (TP/IP)) for syntactic reasons such as Chomsky’s (1982) Extended Projection Principle, according to which all clauses must have grammatical subjects in Spec TP/IP. According to Heim and Kratzer’s (1998) theory of movement, all determiner phrases (DPs), including quantifiers (i.e., expressions of type \(<<e,t>,t>\)) and names (i.e., expressions of type \(e\)), can raise up to a higher position, resulting in two effects (see Heim & Kratzer 1998: ch.8): (i) it leaves a
trace $t$ (= variable) and (ii) it introduces a variable binder (which is interpreted as $\lambda t$) that binds the trace, right below the moved phrase. Since ‘$\pi$’ is neutral between individuals and worlds, it can be bound by the variable binder introduced by the movement of ‘Ann’. So, at this point, the structure of (5) looks like (9).

(9) $[\text{TP/IP Ann } [\lambda t. \text{Ought} \langle f(t), g(t) \rangle \text{ VP } t \text{ is wronged by Betty rather than wrongs her}]]$

(10) $[\text{TP/IP Ought} \langle f(\text{Ann}), g(\text{Ann}) \rangle \text{ VP Ann is wronged by Betty rather than wrongs her}]]$

Since (9) is equivalent to (10) and (10) is a structure for agentive interpretations, (5) does have an agentive interpretation.

On the other hand, on this account, (6) does not have an agentive interpretation. The consideration so far suggests that for an ‘ought’-statement to get an agentive interpretation, the subject should move up to the subject position of ‘ought’. However, in (6a), it is impossible for ‘Ann’ to move up from the subject position of the verb ‘be wronged by Betty rather than wrong her’ to the subject position of ‘ought’, because the subject position of ‘ought’ is already taken up by the expletive pronoun ‘It’. Therefore, (6a) lacks an agentive reading for Ann. The same point applies to (6b).

To summarize, if one is sympathetic with Finlay and Snedegar’s intuitions, my analysis remains unchallenged and so does not need to be revised. If one is sympathetic with Schroeder’s intuitions, they have a good reason to embrace both my analysis of ‘ought’ and Heim and Kratzer’s theory of movement as a package.

---

13 If conversational backgrounds are functions of worlds, that means that a binding theory for world variables needs to be supplied (cf. Chomsky’s Binding Conditions). Percus (2000) and von Fintel & Heim (2020) propose such a binding theory for world variables, according to which a world variable must be co-indexed with (or bound by) the nearest $\lambda$-operator above it. If this binding condition for world variables generalizes to variables over worlds/individuals (or centered worlds), the variables over worlds/individuals must be bound by the nearest $\lambda$-operator above it. For independent motivations for this binding condition on world variables, see Percus 2000 and von Fintel & Heim 2020: ch.5.
4. Conclusion

We have seen Almotahari and Rabern’s puzzle about agentive ‘ought’ and how it raises a significant challenge to the standard theory of ‘ought’. In response to this challenge, I have developed a refinement of the standard theory and showed how it can solve the puzzle. This solution shows at least one way the standard theory can be refined to capture the agent relativity of ‘ought’.14

Seoul National University
Republic of Korea
junhjo.lee@snu.ac.kr

References


14 For helpful conversations and comments on earlier drafts of this paper, I thank WooJin Chung and the audience at the Korean Society for Analytic Philosophy (KSAP) Winter Meeting, 2023. I also thank two anonymous referees for this journal for very thorough and insightful comments.


