A Note on the Good Samaritan Paradox and the Disquotation Theory of Propositional Content[a]

Steven Orla Kimbrough

University of Pennsylvania

kimbrough@wharton.upenn.edu

DEON, May 2002, London

[a]File: deon02-good-samaritan-slides.txt. Revision: 5/21/02.

A word on motivation

Applications in mind. Origin of work in problems of EDI: no compositional semantics $\rightsquigarrow$ expensive kludges.

Big issues: (a) intensionality, (b) getting the inferences right (inclusion, exclusion).

Key elements of my approach: speech acts, F(P); event (subatomic) semantics. Modelling three kinds of logical/semantic structure:

1. Propositional structure. $\approx$ FOL with event semantics
2. Entity structure. $\approx$ FOL with event semantics to model common represenational techniques, e.g., UML, E-R diagrams, trees, networks, ...
3. Intensional structure: speech acts, deontic reasoning, etc. $\approx$ FOL with event semantics plus disquotation theory

Reviewing the Good Samaritan Paradox

Standard deontic logic (SDL) contains the following rule.

Expression 1 *If* $\vdash \phi \to \psi$, *then* $\vdash \mathcal{O}\phi \to \mathcal{O}\psi$

This rule is also known as the good Samaritan Paradox, since it fits (with allowance for stylistic variance) the form:

ϕ: The good Samaritan helps the victim, who has been hurt.

ψ: The victim has been hurt.

But then it follows—given that it is obligatory to help a victim who has been hurt—that it is obligatory that ψ, and surely that is paradoxical, if not absolutely wrong.

Reviewing the Good Samaritan Paradox (con't.)

The Paradox of the Knower, or the Paradox of Epistemic Obligation is a similar, if not identical, problem. Suppose:

1. ϕ
2. $\mathcal{F}\phi$

   (It is forbidden that ϕ.)
3. If ϕ, then $\mathcal{O}\mathcal{K}_i\phi$

   (If ϕ, then [pick your favorite version] it is obligatory that i know that ϕ.)

Now, (1) and (3) imply

4. $\mathcal{O}\mathcal{K}_i\phi$

Add in the epistemic principle

5. $\mathcal{K}_i\phi \rightarrow \phi$

(If i knows that ϕ, then ϕ is true.)

and with Expression 1 you get

6. $\mathcal{O}\phi$

Let ϕ be "i's wife is committing adultery," and we have a forceful example (to many non-swinging husbands at least) of a paradox.

The Anderson reduction is a step in the right direction.

Briefly, instead of $\mathcal{O}\phi$ for "It ought to be the case that ϕ" we have $\Box(\neg\phi \rightarrow V)$ where V is the bad (violation) condition. That is, "ϕ ought to be true" is unpacked as "Necessarily, if ϕ isn't true, then the bad happens" (and that's not good!).

Under the Anderson reduction every violation is equivalent in the sense that V obtains. Let us suppose (temporarily without justification) that instead every deontic condition (obligation, forbiddance, etc.) can be individuated in the sense that each is associated with a possibly unique violation, $V(x)$. Specifically, assume that for every statement ϕ

1. $\mathcal{O}\phi \rightarrow \exists x(\neg\phi \leftrightarrow V(x))$
2. $\mathcal{F}\phi \rightarrow \exists x(\phi \leftrightarrow V(x))$

(I assume that x does not occur freely in ϕ.)

Revisit now the argument

It is instructive to impose existential instantiation where permitted in the argument above. Here in brief is what happens.

1. ϕ
2. $(V(a) \leftrightarrow \phi)$
3. If ϕ, then $(\neg\mathcal{K}_i\phi \leftrightarrow V(b))$
4. $(\neg\mathcal{K}_i\phi \leftrightarrow V(b))$ (From (1) and (3))

Add in the epistemic principle

5. $\mathcal{K}_i\phi \rightarrow \phi$

and with*out* Expression 1 you no longer get

6. $\mathcal{O}\phi$

Instead, you do get

7. $(\neg\phi \rightarrow V(b))$ . . . which I find unparadoxical.

(Some) Points arising:

1. We are distinguishing two obligations: the obligation for spouse a not to commit adultery and the (conditioned) obligation that spouse b (or i) know about the adultery of spouse a. This allows us to make sense of the conclusion, (7). It is simply another case of the paradox of material implication, something most (or many) are happy to live with. If the wife is committing adultery, then if the wife is not committing adultery, then [anything you want]. And note that $(\neg\phi \rightarrow V(a))$ does not follow.[a]

[a] That is, $(\neg\phi \rightarrow V(b))$ follows from $\mathcal{K}_i\phi \rightarrow \phi$ and $\neg\mathcal{K}_i\phi \leftrightarrow (V(b))$, but $(\neg\phi \rightarrow V(a))$ does not. Both follow from ϕ.

(Some) Points arising (con't.):

2. We've done without Expression 1, but something similar obtains.

   Expression 2 *If* $\vdash \phi \rightarrow \psi$, *then*
   $\vdash \exists x(\neg\phi \leftrightarrow V(x)) \rightarrow \exists x(\neg\psi \rightarrow V(x))$

   I want to suggest that the intuition served by Expression 1 is also served—at least nearly as well and without the paradoxical consequences—by Expression 2. But I defer to another time an argument for this suggestion.

3. The approach just outlined is a prototype presented for the sake of exposition. It raises a great number of semantic and even metaphysical questions. I will have a little to say about these issues in the concluding section. First, let us see how the prototype might be embedded in an approach that promises to be a serious candidate.

The disquotation theory

permits us to get (analogs of)

1. $\mathcal{O}\phi \rightarrow \exists x(\neg\phi \leftrightarrow V(x))$
2. $\mathcal{F}\phi \rightarrow \exists x(\phi \leftrightarrow V(x))$

Applies, I believe, to representation of all intensional expressions.

Disquotation theory/approach: core idea

Propositional content has (at least) two important aspects. First, it is about something, that is to say it is true-or-false or rather it is a description, accurate or not, of something. Second, it is itself something about which we attribute certain properties, e.g., that Mary believes it or hopes it or asserts it or promises it. Summarizing (perhaps sloganizing), we might put the point by saying that the sentences of interest here large have the structure: content + comment (on the content). The core idea I wish to develop involves directly recognizing and representing these two aspects (content, comment) of sentences with propositional content.

Disquotation theory: altenative to the modal theory ($\Box$)

Consider the simple propositional content (and speech act) sentence:

Expression 3 *Mary asserts that Sam arrived yesterday.*

My idea is to represent this (and similar) sentence(s) with two kinds of expression: (a) a fundamental expression and (b) one or more axiom schemas, used to articulate meaning for the fundamental expressions. First, we can represent *Sam arrived yesterday* in what is more or less standard event semantics:

Expression 4 $\exists e'(arrive(e') \wedge Subject(e', Sam) \wedge Cul(e', yesterday))$

Let ϕ represent Expression 4.

The fundamental expression for the sentence (in Expression 3) becomes, in shorthand:

Expression 5 $\exists e(assert(e) \wedge Subject(e, Mary) \wedge Obj(e, \ulcorner\phi\urcorner))$

or fully written out:

Expression 6 $\exists e(assert(e) \wedge Subject(e, Mary) \wedge Obj(e, \ulcorner\exists e'(arrive(e') \wedge Subject(e', Sam) \wedge Cul(e', yesterday))\urcorner))$

Thus, the main idea in the fundamental expressions is to treat a quoted sentence (the propositional content) as an object or individual about which a comment is made. In particular, the quoted sentence is the direct object of an event (or eventuality). Moreover, a special form of quotation is used: $\ulcorner\cdot\urcorner$. By quoting an expression in this way—as in 5 and 6—we treat it as an individual and so capture (I argue) the second aspect noted about it.

Disquotation theory: altenative to the modal theory ($\Box$)

Formally we have the following rule:

Axiom Schema 1 (Assert Rule) $\forall e((assert(e) \wedge Obj(e, \ulcorner\phi\urcorner)) \rightarrow (Veridical(e) \leftrightarrow \phi))$

Axiom Schema 1 should be thought of as a rule into which we may substitute uniformly for ϕ any well-formed formula in the current language.

Note: This generalizes to all the speech acts.

Disquotation theory: Extension to deontic reasoning

Obligations and permissions. Problems with the standard logic. Now an alternative.

The Anderson reduction: Instead of $\mathcal{O}\phi$ for "It ought to be the case that ϕ" we have $\Box(\neg\phi \rightarrow V)$ where V is the bad (violation) condition. That is, "ϕ ought to be true" is unpacked as "Necessarily, if ϕ isn't true, then the bad happens" (and that's not good!).

Then. . . Suppose that a delivery is obligated:

Expression 7 $\mathcal{O}\exists e_1(deliver(e_1) \wedge Sub(e_1,\ a) \wedge Obj(e_1,g) \wedge IndObj(e_1,\ s) \wedge Sake(e_1,\ e))$

Disquotation theory: Deontic reasoning

Our fundamental schema for *ought* follows the usual form

Fundamental Schema 1 (Ought) $\exists e(ought(e) \wedge Obj(e, \ulcorner \phi \urcorner) \wedge \Gamma)$

and our example (Expression 7) instantiates in the predictable fashion:

Expression 8 $\exists e(ought(e) \wedge Obj(e, \ulcorner \exists e_1(deliver(e_1) \wedge Sub(e_1, a) \wedge Obj(e_1, g) \wedge IndObj(e_1, s) \wedge Sake(e_1, \boxed{e}) \urcorner))$

Corresponding closely to the spirit of the Anderson reduction gives us the *weak* ought rule:

Axiom Schema 2 (Weak Ought Rule) $\forall e((ought(e) \wedge Obj(e, \ulcorner \phi \urcorner)) \rightarrow (\neg\phi \rightarrow V(e)))$

Disquotation theory: Deontic reasoning

My theory allows us to employ the *strong* ought rule:

Axiom Schema 3 (Strong Ought Rule)
$\forall e((ought(e) \wedge Obj(e, \ulcorner\phi\urcorner)) \rightarrow (\neg\phi \leftrightarrow V(e)))$

Permission works similarly.

Expression 9 (Permission) $\exists e(permit(e) \wedge Obj(e,$
$\ulcorner \exists e_1(deliver(e_1) \wedge Sub(e_1, a) \wedge Obj(e_1, g) \wedge IndObj(e_1, s) \wedge$
$Sake(e_1, \boxed{e}) \urcorner))$

Disquotation theory: Deontic reasoning, directed obligation

The move here is the same as in systems of obligation and permission: add predicates to qualify the underlying event. If a has an obligation to b under system of norms n that ϕ, we then have:

Fundamental Schema 2 (Directed Obligation) $\exists e(ought(e) \wedge Subject(e,\ a) \wedge IndObj(e,\ b) \wedge IsUnderNSystem(e,\ n) \wedge Obj(e,\ \lceil \phi \rceil) \wedge \Gamma)$

Note: Instead of $V(e)$ we need $V(e,n)$.

Key points

1. Fundamental schema: form for representing the basic intensional expression.

2. The quotation operator affords the highest possible degree of intensionality.

3. Axiom schemas: let you relax the intensionalty of the quotation operator.

4. The approach affords reification of obligations, necessities, speech acts, etc., which accords with ordinary usage.

FAQs

- Where do the xs (eventualities: events, states, processes, etc.) in the $V(x)$s (etc.) come from? And what to they refer to?

  These questions call for a philosophical account of the origins of obligations, permissions, and so on. The disquotation theory should be neutral with respect to any reasonable philosophical theory.

  A better question is:

- Does the disquotation theory help us in any way to develop a naturalistic account of deontic notions?

  Maybe. By convention, certain doings are recognized as speech acts, which may be referred to by eventuality variables/constants. Necessities may arise for a variety of reasons. Permissions and obligations? Compatible with a conventionalist view (with institutional powers).

FAQs (con't.)

- What about other paradoxes of deontic logic and reasoning, specifically contrary-to-duty obligations?

  I conjecture that the theory works, or at least helps, here too. As we have seen, it allows distinction between individual (kinds of?) obligations. There need be no paradox in contrary-to-that-duty obligations. Killing and killing mercifully may both be forbidden, but with distinct forbiddances, and hence violated separately. Think of "throwing the book" as listing the distinct violations.

  But here, and elsewhere, lie opportunities for further research.