Dynamic Expressivism about Deontic Modality (*Abstract*)

William Starr, Department of Philosophy, Cornell University
will.starr@cornell.edu :: http://williamstarr.net

This paper presents an expressivist semantics for deontic *may* and *must* using the resources of dynamic semantics. It is shown that the Frege-Geach problem does not afflict this semantics. I conclude by discussing the scope of this semantics and its relevance to the philosophical project of metaethical expressivism.[1]

In the dynamic semantics presented here, the meaning of a sentence is the characteristic role that sentence plays in changing agents' mental states. Purely informational attitudes (e.g. belief) are modeled in terms of their content: the way they represent the world to be. The set of possible worlds that are that way provides a simple formalization of such an attitude. The elimination of worlds from that set can then provide a model of the way an informational attitude changes when more information is accepted. By contrast, conative attitudes (e.g. desire) are not here modeled (just) in terms of content: they value certain ways the world could be over others. Formally, this can be modeled as an ordering of possible worlds. The meaning of some sentences, like *Sally swam*, can be completely captured by specifying how they change purely informational attitudes, i.e. in how they eliminate worlds. Sentences like *Sally must swim* convey that certain ways the world could be are valued over others. Can this be correctly modeled as the elimination of worlds? If it were, it would amount to representing the *world* (or an agent in that world) as making certain possibilities more valuable than others. In other words, it would describe a matter of fact about the value of swimming. On this **descriptivist** approach, deontic communication proceeds by aligning conative attitudes with the facts about values. However, one might instead model deontic statements as attempts to coordinate conative attitudes without describing facts about value. This amounts to coordinating on an ordering by means other than eliminating worlds where that ordering lacks the target properties, i.e. providing information *about* the ordering. Any **expressivist** approach like this must respond to the Frege-Geach problem (Frege 1923; Geach 1965; Schroeder 2008): how can expressive vocabulary meaningfully embed under truth-conditional operators like negation and *if*? After providing a formalization of an expressivist semantics for *may* and *must*, I will show that this problem can be solved. The standard meanings for negation and conditionals in dynamic semantics capture this kind of embedding while predicting that these operators behave classically when non-modal vocabulary are involved. This is not a full vindication of metaethical expressivism; it covers a quite limited deontic vocabulary in one language, and it remains to be shown that this semantics adheres to the exact philosophical doctrines needed in metaethics. But it does provide an instructive first step and allows one to more precisely examine the relation between an expressivist formal semantics and the expressivist project in metaethics.

Suppose we choose to view meaning from the dynamic perspective, identifying it with the characteristic role a sentence plays in linguistic activities (conversation, calculation, inquiry, thought, etc.). How should sentence meanings be formally specified? On the Stalnakerian model of these activities – which I shall modify shortly – their progress is

[1] A similar approach is suggested by Willer (2010: §5.5). The present work provides a more thorough expressivist interpretation of dynamic semantics and a different approach to deontic operators.

modeled by an evolving body of contextual information, c. So the characteristic role of a sentence is to change an arbitrary state of information c into another c'. This can be modeled as a function from c to c' and specified by an equation of the form $c[\phi] = c'$, which says just how ϕ changes c.[2] (Read '$c[\phi]$' as 'c updated with ϕ'.) For illustration, consider a standard propositional language with atomic sentences $\mathcal{A}t = \{\mathsf{p}_0, \mathsf{p}_1, \ldots\}$. A possible world is treated as a valuation, i.e. an assignment of either 1 (True) or 0 (False), to each atomic sentence.

Definition 1 (Update Semantics)

(1)	$c[\mathsf{p}] = \{w \in c \mid w(\mathsf{p}) = 1\}$	(2)	$c[\neg\phi] = c - c[\phi]$
(3)	$c[\phi \wedge \psi] = (c[\phi])[\psi]$	(4)	$c[\phi \vee \psi] = c[\phi] \cup c[\psi]$

While these clauses *look* different from a standard possible worlds semantics, what exactly is the difference and what does it offer? There are some ways of interacting with c that are not, intuitively, descriptive: they do not amount to providing information (eliminating worlds lacking some feature). None of the above operators interact with c in this non-descriptive way, and until one adds operators that do, the clauses of Definition 1 are equivalent to the standard ones. The difference between expressive and descriptive operators is reflected in the key logical concept of dynamic semantics, **support**, and its relation to the key logical concept of classical semantics, **truth**.

Definition 2 (Support, Truth in w)

(1) Support $c \vDash \phi \Leftrightarrow c[\phi] = c$ (2) Truth in w $w \vDash \phi \Leftrightarrow \{w\}[\phi] = \{w\}$

When the information provided by a sentence is already contained in c (explicitly, or implicitly), c is said to support that sentence. Truth is a special case of support (Starr 2010). Being true at a world amounts to being supported by perfect information about that world. Truth tracks the constraints a sentence places on a world, i.e. how it describes a world. Support tracks the constraints a sentence places on a state of information, and therefore on informational attitudes. As this suggests, the two concepts come apart when an operator exploits the sole difference between information states and worlds: uncertainty. Veltman's (1996) simple semantics for the epistemic modal *might* illustrates this: $c[\mathsf{Might(p)}] = \{w \in c \mid c[\mathsf{p}] \neq \emptyset\}$. On this semantics, Might(p) does not place a constraint on w, but rather on c. Compare a semantics in the spirit of modal logic, using $R(w)$ to mean the worlds accessible from w: $c[\mathsf{Might(p)}] = \{w \in c \mid R(w)[\mathsf{p}] \neq \emptyset\}$. This operator provides information about the facts in worlds, namely the modal facts captured with R. As a result of this difference, the two sentences interact differently with c. While the latter, descriptivist, semantics focuses on a world, and eliminates it from c if the modal facts at that world aren't right, the former semantics tests c for compatibility with p. The former falls into the familiar descriptivist model of communication: worlds are eliminated from consideration when they don't have the right properties. The latter fits the expressivist model: it places a constraint on the agent's attitudes without expressing a proposition about those attitudes. While this perspective on Veltman's semantics is new, it is merely the inspiration for the expressivist approach to deontic modals I shall now outline.

While epistemic modals express constraints on informational (cognitive) attitudes, de-

[2] $[\phi]$ is a function which applies to c and delivers c'. The notation $[\phi](c)$ seems more appropriate, but I will follow precedent. (Veltman 1996)

ontic modals express constraints on evaluative (conative) attitudes. Decision theory has accustomed us to and illustrated the power of viewing **preference** as a key conative attitude in practical reasoning. Formally, preferences will be modeled as binary orderings on worlds: $w_1 \succeq w_0$, read 'w_1 is at least as preferable as w_0'. When $w_0 \succeq w_1$ and $w_1 \succeq w_0$, the agents are indifferent about w_0 and w_1 ($w_0 \equiv w_1$). When $w_0 \succeq w_1$ and $w_1 \not\succeq w_0$, the agents strictly prefer w_0 to w_1 ($w_0 \succ w_1$). As Hansson (1990) discusses, using preferences in deontic logic gives the ordering an intuitive and theoretically cohesive interpretation, and offers insight on some of the paradoxes facing standard deontic logic. My proposal is to combine this with a dynamic semantics: $\mathsf{May}(\mathsf{p})$ and $\mathsf{Must}(\mathsf{p})$ express constraints on preferences, much as $\mathsf{Might}(\mathsf{p})$ expresses a constraint on the informational attitudes that define c. In particular, $\mathsf{Must}(\mathsf{p})$ tests that every live (in c) p-world is strictly preferred to every live $\neg\mathsf{p}$-world. $\mathsf{May}(\mathsf{p})$ tests that some live p-world is strictly preferred to some live $\neg\mathsf{p}$-world. The focus on 'live' worlds reflects that *may* and *must* also have epistemic readings, and even on their deontic readings they are used to communicate about the value of live possibilities (Ninan 2005). On this model, deontic sentences interact with both information and preferences, so updates are defined on $\langle c, \succeq \rangle$, rather than c. I adopt the convention of writing $c_{\succeq}$ instead of $\langle c, \succeq \rangle$ and reading set-theoretic symbols as operating on c rather than the pair, e.g. $c_{\succeq} - c'_{\succeq} = \langle c - c', \succeq \rangle$.[3]

Definition 3 (May and Must)
(1) $c_{\succeq}[\mathsf{May}(\mathsf{p})] = \{w \in c \mid \exists w_0 \in c_{\succeq}[\mathsf{p}] \;\&\; \exists w_1 \in c_{\succeq} - c_{\succeq}[\mathsf{p}] : w_0 \succ w_1\}_{\succeq}$
(2) $c_{\succeq}[\mathsf{Must}(\mathsf{p})] = \{w \in c \mid \forall w_0 \in c_{\succeq}[\mathsf{p}] \;\&\; \forall w_1 \in c_{\succeq} - c_{\succeq}[\mathsf{p}] : w_0 \succ w_1\}_{\succeq}$

On this view, an agent X who issues a deontic modal signals that $c_{\succeq}$ has the right properties to support it. What if the shared understanding of $c_{\succeq}$ does not support it? If X is an authority, c or $\succeq$ changes to support the claim. Otherwise, infelicity and clarification ensue. Using the notational convention for '$c_{\succeq}$', the clauses in Definition 1 extend to a language containing deontic modals. So these deontic modals embed freely under 'truth-conditional' operators. Conditionals can be treated in a number of ways, as material conditionals, as Stalnakerian conditionals, or as a test that c supports the consequent when it's updated with the antecedent: $c_{\succeq}[(\mathsf{if}\,\mathsf{p})\,\mathsf{q}] = \{w \in c \mid c_{\succeq}[\mathsf{p}] \vDash \mathsf{q}\}$. On any of these treatments, deontic modals will be coherently embeddable in antecedents, despite receiving an expressivist semantics. This has been achieved by making all of the language a bit more expressivist, but as noted the semantics for connectives remains classical when limited to a fragment without expressivist operators. Because support is the key concept of this semantics, consequence is defined differently:

Definition 4 (Consequence) $\phi_1, \ldots, \phi_n \vDash \psi \Leftrightarrow \forall c : c[\phi_1] \cdots [\phi_n] \vDash \psi$

This account of consequence also offers an approach to the Miner's, Chisholm's and Gentle Murder paradoxes that does not sacrifice modus ponens (Willer 2012). The focus on support, rather than truth, does not mean that deontic modal sentences do not *have* truth-conditions. By Definition 2.2, they have theoretically useless truth-conditions: $\mathsf{May}(\mathsf{p})$ is false at every world and $\mathsf{Must}(\mathsf{p})$ is true at every world. But this is as it should be, it is the support-conditions that are theoretically central: support conditions do all the work in explaining how deontic modals work in communication and reasoning.

[3]What about operators that *change* $\succeq$? See the semantics for imperatives explored in Starr (2012).

References

Frege, G. (1923). 'Logische Untersuchungen'. *Beiträge zur Philosophie des deutschen Idealismus* **3**: pp. 36–51. References to English translation Frege (1963).

Frege, G. (1963). 'Compound Thoughts'. *Mind* **72**(285): pp. 1-17. Translation of Frege (1923)., URL http://www.jstor.org/stable/2251920.

Geach, P. T. (1965). 'Assertion'. *The Philosophical Review* **74**(4): pp. 449–465. URL http://www.jstor.org/stable/2183123.

Hansson, S. O. (1990). 'Preference-Based Deontic Logic (PDL)'. *Journal of Philosophical Logic* **19**(1): pp. 75–93. URL http://dx.doi.org/10.1007/BF00211186.

Ninan, D. (2005). 'Two Puzzles About Deontic Necessity'. In J. Gajewski, B. N., V. Hacquard & Yalcin, S., eds., 'New Work on Modality, MIT Working Papers in Linguistics', vol. 51. Cambridge, MA: MIT Press. URL http://semanticsarchive.net/Archive/WZlNTg0Y/NinanDeontic.pdf.

Schroeder, M. (2008). 'What is the Frege-Geach Problem?' *Philosophy Compass* **3**(4): pp. 703-720.

Starr, W. B. (2010). *Conditionals, Meaning and Mood*. Ph.D. thesis, Rutgers University, New Brunswick, NJ. URL http://hdl.rutgers.edu/1782.1/rucore10001600001.ETD.000056780.

Starr, W. B. (2012). 'A Preference Semantics for Imperatives'. Ms. Cornell University, URL http://williamstarr.net/research/a_preference_semantics_for_imperatives.pdf.

Veltman, F. (1996). 'Defaults in Update Semantics'. *Journal of Philosophical Logic* **25**(3): pp. 221-261. URL http://dx.doi.org/10.1007/BF00248150.

Willer, M. (2010). *Modality in Flux*. Ph.D. thesis, University of Texas at Austin, Austin, TX. URL http://www.maltewiller.net/Modality_in_Flux.pdf.

Willer, M. (2012). 'Dynamic Thoughts on Ifs and Oughts'. Ms. University of Chicago.