Comparative evaluative judgments

Andrés Soria Ruiz

August 4, 2017

Broadly expressivist proposals about normative language understand value judgments in *binary* terms, that is, in terms of the expression, on the part of the utterer, of a favorable or unfavorable attitude (sometimes called a *PRO*- or *CON*-attitude) towards the object under evaluation. For theories of this sort, when a speaker utters (1) she expresses a favorable attitude towards volunteering for a charity; and when she utters (2), she expresses an unfavorable attitude towards donating money to a charity.

(1) Volunteering for a charity is good.

(2) Donating money to a charity is bad.

This approach faces a fundamental shortcoming when faced with sentences like the following:

(3) Volunteering for a charity is better than donating money.

When a speaker utters (3), she need not endorse and/or reject neither volunteering nor donating. She is merely comparing the *goodness* of the two actions; and her uttering (3) is compatible with adopting almost any combination of positive and negative attitudes towards either of them (with the exception of being in favor of donating money while being against volunteering). This is shown by the fact that (4)-(6) are acceptable, while (7) is not:

(4) Volunteering for a charity is better than donating money, though both are bad.

(5) Volunteering for a charity is better than donating money, though both are good.

(6) Volunteering for a charity is better than donating money; in fact, volunteering for a charity is good whereas donating money is bad.

(7) ?? Volunteering for a charity is better than donating money; in fact, volunteering for a charity is bad whereas donating money is good[1].

[1]That these combinations are coherent suggests that *good* is a relative adjective, in the sense of Kennedy 2007. Other value adjectives show different patterns of inference. In particular, comparisons using negative value adjectives like *bad* or *ugly* invite the inference that the positive form applies to one or both of the *relata*. An anonymous referee points to the case of *beautiful*, where in order to cancel the inference to

How can the expressivist insight about *absolute* judgments of value (i.e. (1) and (2)) be extended to comparative judgments like (3)?

Value adjectives are gradable, so the literature on gradability in semantics should point to a solution. However, value adjectives are different from run-of-the-mill gradable adjectives—adjectives like *tall* or *rich*. In those cases, it is clear what it means for an object a to possess the relevant property to a higher degree than another object b: it is simply to possess more height, or money. But what about evaluative properties? What is for an object to possess an evaluative property, e.g. *goodness*, to a higher degree than another object? This is what we need to spell out. What we propose is to combine insights from the literature on gradability and meta-ethics to arrive at a model that makes the right predictions both for absolute and comparative value judgments.

As basic elements in our semantics we use Gibbard's (1990, 2003) *hyperplans*. Hyperplans were devised by Gibbard as tools for modeling the close connection between normative judgments and action-planning. In his view, to judge that an action is rational is to adopt a plan to perform that action in the appropriate circumstances. A domain H of hyperplans looks very much like the familiar domain W of possible worlds of intensional semantics (i.e. maximally determined states of affairs), and given the usefulness of understanding informational content in terms of set-theoretical operations over W, it is suggestive to understand *normative* content in terms of set-theoretical operations over H (see Field 2009; Yalcin 2017 for suggestions in this direction).

A hyperplan is a maximally decided planning state: a state that tells you what to do in every conceivable situation that you could find yourself in[2]. We can think of a hyperplan as a *total* function from the set of conceivable situations S to the set of possible actions A. The actual plans adopted by agents however, are less than maximally decided: for many situations, they do not tell you what to do. We can thus conceive of a plan as *partial* function from S to A, or alternatively, as a set of hyperplans that agree on what to do in some situations, but not for others. Conversely, an action a can be defined as the set of hyperplan-situation pairs $\langle h, s\rangle$ such that the agent of h performs action a in s.

Plans and situations can be now employed to give truth-conditions for absolute judgments of value. Following the expressivist tradition, we map the adjective *good* (*bad*) at a context of utterance c to a relation of *support* (*rejection*) by the relevant plan P at c.

We start by defining support and rejection (at a situation s) as follows:

P supports a in s iff $\forall h \in P.\langle h, s\rangle \in a$

P rejects a in s iff $\forall h \in P.\langle h, s\rangle \notin a$

---

ascribing the positive form to either *relatum* some qualifying particle is needed:

(1) Anna is more beautiful than Berta, ?? but neither of them is beautiful.

(2) Anna is more beautiful than Berta, but *in fact* neither of them is beautiful.

[2]Every hyperplan will tell what to do if your car breaks, if it doesn't, if there's a fire, if your neighbors fight, if you were Ceasar right before crossing the Rubicon, etc.

That is, P supports (rejects) a in a situation s just in case every hyperplan in P is such that the agent of h does (does not) a in s. If neither condition holds, then P is *indifferent* with respect to a.

Now as we've seen, plans are defined for more than one situation, so in order to generalize the notion of support/rejection we need to consider a set of situations, with some restrictions. Intuitively, we want to say that a plan supports, say, smoking, just in case most normal situations in which one could smoke are situations in which one actually smokes. Let us stipulate that, for every action a, there exists a set of a-pertinent situations, that we will loosely define as situations where action a could be performed (think of them as situations in which nothing prevents you from performing action a). Now our definitions of support and rejection by a plan can be generalized to pertinent situations as follows:

> P supports a iff $\forall h \in P$ & $\forall s$ s.t. s is a-pertinent, $\langle h, s\rangle \in a$ (*ceteris paribus*)
>
> P rejects a iff $\forall h \in P$ & $\forall s$ s.t. s is a-pertinent, $\langle h, s\rangle \notin a$ (*ceteris paribus*)

In words: P supports (rejects) a just in case every hyperplan h in P and a-pertinent situation s are such that the agent of h does (does not) a in s, *ceteris paribus*[3].

Now, in order to derive comparisons from this system, we need to restrict the set of situations that we are considering in a different way. In particular, for any two actions a and b, we need to consider only situations that are a- and b-pertinent. Such restriction on our original plan P delivers a set of "subplans" of P, and all we have to do is consider which of a or b is good (or bad) relative to those subplans:

> $a >_P b$ iff $\forall P' \subseteq P$. and $\forall s \in P'$ s.t. s is both a- and b-pertinent, P' supports a and rejects b.

That is, an action a is better than an action b relative to a plan P just in case, given a choice between a and b, we would consistently choose a over b without modifying our plan, that is, without adopting a different set of hyperplans.

Informally, the idea is that a plan may be such that any number of actions is supported and rejected by it in different situations, but in order to make a comparative judgment, it doesn't matter whether the actions are actually supported or rejected. All that matters is that, having to choose, we would choose one over the other. This predicts the admissibility of (4)-(6). Nonetheless, the proposed truth conditions do rule out a situation like (7), where volunteering is better than donating, and yet donating is supported while volunteering is rejected: if we adopt a plan such that every volunteering-pertinent situation is one where we don't volunteer, and every donating-pertinent situation is one where we do donate, then

[3]This *ceteris paribus* clause is meant to help with the following, immediate problem: for many actions a and b, there will be situations that are both a- and b-pertinent, but where both actions cannot be performed. For instance, I may support jogging and smoking, yet reject doing both at the same time. By our naked definition however, if I end up jogging in a smoking-pertinent situation, then smoking comes out "not good" relative to my plan. The *ceteris paribus* clause is meant to read as: assuming that nothing else is supported by this plan.

if there's any situation that is both vounteering- and donating-pertinent, we will always donate rather than volunteer.

Importantly, this strategy preserves the syntactic primacy of the absolute over the comparative form of value adjectives (see Barker 2002; Benthem 1982; Klein 1980). The expressivist idea that absolute value judgments express outright support or rejection of an action is the starting point from which a semantics for the comparative form is derived.

Finally, disagreement over a claim like (3) reveals different ordering preferences between the plans adopted by the disagreeing speakers. But the proposed semantics does not, by itself, provide an account of disagreement. That depends on exactly how we define the content expressed by the claims involved. For example, a subjectivist could adopt our system and say that a claim like (3) and its negation *describes* the plans of their utterers. On this view, the two speakers would be talking past each other. In order to hold on to a semantics where different speakers may adopt different plans and retain a notion of disagreement, we may recur to the expressivist idea that a normative disagreement is a practical disagreement, that is, not a disagreement about what follows from a certain plan, but about what plan to adopt[4].

References

Barker, Chris (2002). "The dynamics of vagueness". In: *Linguistics and Philosophy* 25.1, pp. 1–36. ISSN: 0165-0157. DOI: 10.1353/pew.2002.0034.

Benthem, J Van (1982). "Later than late: On the logical origin of the temporal order". In: *Pacific Philosophical Quarterly anc Personalist (The)*.

Field, Hartry (2009). "Epistemology without metaphysics". In: *Philosophical Studies* 143.2, pp. 249–290. ISSN: 00318116. DOI: 10.1007/s11098-009-9338-1. arXiv: arXiv:1011.1669v3.

Gibbard, Allan (1990). *Wise Choices, Apt Feelings*. Clarendon Press ; Oxford University Press.

— (2003). *Thinking how to live*. Harvard University Press. ISBN: 0674037588.

Gust, Helmar and Carla Umbach (2015). "Making use of similarity in referential semantics". In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Vol. 9405, pp. 425–439. ISBN: 9783319255903. DOI: 10.1007/978-3-319-25591-0_31.

Kennedy, Christopher (2007). "Vagueness and grammar: The semantics of relative and absolute gradable adjectives". In: *Linguistics and Philosophy* 30.1, pp. 1–45. ISSN: 01650157. DOI: 10.1007/s10988-006-9008-0.

[4] An anonymous reviewer suggests a parallelism between this system and Gust and Umbach 2015's proposal to treat aesthetic adjectives using multidimensional spaces. Reasons of space prevent me from exploring this connection, but it will be taken into account in the presentation.

Klein, Ewan (1980). *A semantics for positive and comparative adjectives*. DOI: 10.1007/BF00351812.
Yalcin, Seth (2017). "Expressivism by force". In: *New Work on Speech Acts*.