*Reconstructing something we discussed the week before last week by memory.*

A person S values X =

**S believes that X is good (instrumentally or intrinsically) *and* S desires X *and* S treats the desires for X as a justification for obtaining X.**

This analysis of valuing, which is defended by the philosopher Michael Bratman, suggests that our algorithm will need to do three things:

***Step 1.*** “Inventing” categories of value (i.e., aesthetic value for our purpose) — by unsupervised learning?

This requires tracking people’s desire patterns (Bratman’s analysis implies that, although one can desire what one does *not* value, if one values something, one must desire it.

Then, we need to find certain indicators to separate desires that people treat *as justification* and desires that are not treated as justification for purchase (a possibility we considered: based on reviews that people are willing to go public about). Bratman’s analysis of value says that only desires that are considered justification for acting on the desired count as value.

Use unsupervised learning to find patters in people’s values in books.

***Step 2.*** For a particular user, try to fit that person into one of the categories of value we obtained from Step 1.

***Step 3.*** Based on the paradigm pattern of people in that category, recommend next book(s) to read.

I take it that our task is to make informed guesses about which category of valuing pattern — based on Step 1 — the user *WOULD* develop (i.e., what kind of aesthetic value they have) ***if*** our algorithm did not intervene and the user were just left to make his/her own choices given perfect information.

*P.S. The ethics of paternalism paper I sent you in summer argues that information about a person’s will/value does not yet tell us immediately what we should do and that depends on whether the person is an expert or not. I am not yet sure what to make of it in the context of our work.*