Sajid Sadi

Week 3

Handbook of Personal Relationships: Theory, research, and interventions, chapter "Self-Report Methods in Studying Personal Relationships"

The quotation "…we seek to understand others' imputation of meaning for what occurs in their lives" particularly resonated with me in this matter, because it highlights the fact that relationships are in fact a subjective experience. It brings to mind the previous in-class discussion regarding whether the stereotypes that people have about what machines are and what they can do will stand in the way of relationships. As it stands, machines (even those programmed to be relational) have a relatively limited understanding of the meaning of the relationship within their own "lives". This leads me to think: perhaps machines need to be more aware of the effect of relationships on themselves in order to be able to have meaningful relationships. Perhaps, if the relationships had meaning to the machines, then they would be more meaningful to humans as well. Of course, this leads us down the long path to building a Turing Test passing machine (and it would have to be exactly that advanced in order to hold a relationship… it need not be actually intelligent), but I think it still deserves some thought, simply in order to scope the extents to which we can build a relational machine without it being quite that advanced.

On a different note, as the authors expound over many words, self-reporting seems to be the best window into relationships, simply because relationships include so much "hidden state" that we otherwise have no access to. It is probably also a very good way of finding out about the transition in a person from accepting a machine as a machine to accepting a machine as a partner in a relationship. However, the issues of the test interference are probably much worse for machines, since there are so many more stereotypes and socially-undefined conditions within such a relationship, which in turn leads to a high level of data tainting or biasing. A post-fact interview may solve some of these issues, but in turn would obscure the most interesting mental state information, which is generally "overwritten" in memory as the relationship progresses. Adding generality to bypass the social issues would also have a similar negative effect, but due to loss of specificity, which confounds the different factors. Until these issues are resolved, either in terms of social acceptance (chicken and egg problem here) or in terms of mathematically separating out the tainting, I remain somewhat skeptical of comparing human-machine data of this sort to human-human data.

On a final positive note, it would be interesting if the machine kept the diary instead of the user. It would, in fact, be another use for a relational machine (ignoring, for the second, all other issues), and might provide an interesting look into how a relationship works for humans.

Handbook of Personal Relationships: Theory, Research, and Interventions, chapter "The Analysis of Data from Two-Person Relationships"

This paper, and the one after it, provides a very nice technical overview of how data of this sort is analyzed, though I must admit that the jargon lies heavy upon the pages. The authors go on at length to describe the issues with not considering a dyad independent in many cases. However, an interesting situation arises when one member of a dyad is a machine, since the machine's "thoughts" can be logged quite completely. In such a case, perhaps it is theoretically possible to consider the two individuals separately (or at least, deconstruct and discount the effects on the machine upon the relationship). Of course, even knowing the machine's thoughts may be very insufficient in mapping out what the impact on the relationship itself was, since the machine's thoughts would most likely be different from human ones.

On a mathematical aside, the authors point out that synthesizing indices from sub-aspects which may themselves be interesting causes index confusion. However, success in relationships is in itself such a composite index, and is composed of aspects and self-disclosure, trust, bond, etc. It serves as a good reminder that often times it is important to break down the human-machine relationship and consider just the aspects for which a machine was designed (or vice versa, to design a machine towards those aspects).

It also strikes me that a lot of the techniques that are discussed are the same ones via which machines themselves learn. For example, the log linear approach to sociometry is similar to approaches used in data mining from unfiltered sources and deemphasizing baseline noise in the system. Time-domain modeling seems reminiscent of support vector modeling, and the Markov model is of course the counterpart to HMMs, which are used to generically learn any and everything. On one hand, it points out that machine learning is very much parallel to the sort of quests these papers chronicle. However, it also brings out the question of whether these methods are themselves valid in assessing a system that uses similar methods to generate its behavior. If they are not valid, then perhaps a new set of methods is needed for assessing such relationship that are not affected by the artifacts of this sort of determinism.

The Measurement of Trust and its Relationship to Self-Disclosure

When I began reading this paper, my very first question was: does the idea of self-disclosure to a machine change the dynamics of trust? That is, what changes when the recipient of the information is a machine that does not inherently understand the idea of trust? When we speak of trustworthy computing, we really speak of trustworthy policy, and not of trust as a primitive. Moreover, even if the machine claimed to understand trust (via policy or otherwise), it can in fact be made to recount that information easily, since its innards are obviously digital and thus accessible at least to its creator and any other unscrupulous or subpoena–bearing individual. This leads to a new form of temporally and individually distributed trust that has hereto not existed.

The top of page 252 makes some interesting points:

1. Trust is needed in self disclosure –this is almost so by definition
2. People self-disclose to those whom they don not sufficiently distrust

3. One often does not self disclose to trusted parties – trust is an insufficient condition for it

This suggests that there is a high level of volitional/introspective control imposed upon divulgence, which may stand against machine relationships due to issues regarding humans being societally programmed, as it stands, to automatically box machines under the "not alive/conversational/relational" category. This is also supported by the conclusion that disclosure did not ramp up with individual trust, indicating that even if the machine did everything right, a social realignment would be necessary nonetheless. Unfortunately, self-disclosure is more important to machines due to their limited ability to infer from context, first pass impressions, etc., than it is for human beings. On the brighter side, however, relationships of the "human-to-expert system" type face less of these restrictions, and in those cases the lack of self-disclosure may very much be the "normal" condition.

Development and Validation of the Working Alliance Inventory

In light of the other readings, it struck me as immediately strange in reading the abstract that the authors had considered the three components of a working alliance (bonds, goals, tasks) as statistically independent in a social setting. It seemed far more likely given the complexity of human relationships that the three parts would be inherently statistically interdependent. Especially, it seems somewhat ludicrous to believe that bonds do not have an effect on collective goals. Additionally, I am not entirely convinced that the artificial attempts to separate the components do not explain to some extent the outcomes of the study into the separation, but that does not strike me as very important to the goals of the course.

Bordin's theory seems to me to be a good approach to human-machine relationship insofar as cooperation and concordance within a specific goal or set of goals is desirable. In particular, the fact that he presents his working alliance as a vehicle for interaction simplifies its mapping to such relationships. It also bypasses some of the thorny issues that surround non-working relations, which of course tend to pain the machine as more anthropomorphic and "alive" than a purely working relationship. Bordin also asserts that the balance of the 3 components is fluid, which brings to mind Duck's assertion that relationships require different components over time. It might be interesting if the machine could try to map out some of these factors (ie, take part of it test for itself using its own internal state regarding a relationship) in order to find out where it stands in a relationship and therefore adapt accordingly, especially since the three components seem to be information that it either must extract to function (goals, tasks), or can interpolate from externalized behavior of the other party (bond) to some extent.