Relational Machines

Project Proposal

Sajid   Sadi

Goals & Overview

The goal of this project is to provide the sponsors visiting the Media Lab during the upcoming TTT event in early May with concise and relevant recommendations about projects and groups that would be of interest to them at the point of decision. This is a direct and calculated departure from how things have been done in the past. Previously, sponsors have been provided with a simple listing of groups by keywords, along with a number of "trails" they may follow if their general interests match one of the 3–4 trails offered. The keywords have customarily been highly individual and tied to the groups, thus minimizing the chances that sponsors would actually know or understand what those words represented. The project has three specific phases. In the first phase, data will be collected from the Media Lab faculty and students about the interest of the group. This will be done in a way that promotes a level of convergence and allows inference based on the input. The second phase involves collecting preference information from users. The third phase involves reasoning about the input from the first two phases, as well as user actions, in order to provide the recommendations to the users. Each of these phases requires a different model of user interaction.

The first phase, data collection from Media Lab members, requires the system to be somewhat strict and draconian. Studying previous literature handed out to the sponsors clearly indicate that the freeform selection of keywords leads to problems. Therefore, this interface will directly and indirectly remind the user to conform to guidelines while specifying keywords in order to produce an ontology that is concise, context-free (ie, one does not have to be a member of the lab in order to understand the usage of the term), and associative (ie, acknowledging the fact that all groups have some level of overlap with others, though each does have its own view of the problem spaces and techniques). This is also the portion least sensitive to relational issues, since in some sense I am able to twist arms (by pointing out that bad keyword selections, or worse, lack of selection) will greatly reduce or even null the possibility of the system guiding sponsors in their direction. Nonetheless, the system should be highly accessible in order to minimize the barrier to proper use. However, I will be focusing hereon primarily on the relationship with the rest of the system.

The second phase involves customization, and the primary goal is to gain a concise idea of what the user in interested in. The key factor here is that the user will generally have minimal attention spans, especially if they are using the kiosks at the registration area on the day of the meet. Therefore, the process should be novel enough to capture attention for the necessary 1–2 minutes. It must also avoid the pitfall of appearing boring or useless (particularly, any sort of a form must be avoided), because this will hurt adoption prima fascia, even before the user has had a chance to find out why it may be useful. While one is more than welcome to go back and re-enter preferences, the TTT schedule will generally prevent that, and the system must make the best of what it can get. Here the relationship prototype is a quick chat with an interesting person who is interested in you. The interface must focus directly

on the user, and must at all times make obvious that this is all about the user's experience, not some project the user may not even be interested in.

The third phase is the point-of-decision recommendation system. Here, the relationship model has changed minimally from my previous expectations. I expect the system to appear more like an experienced butler or a good real estate agent, rather than a personal assistant. The latter focus too greatly on a one-on-one metaphor, which is somewhat difficult to sustain via the kiosks through which the users will interact with the system. Additionally, it is very clearly a public system, and the relationship model needs to reflect that professionalism and detachment. At the same time, the idea of service comes into play. The system must be of service to the user, and must be graceful in face of a failure. Since interaction is only initiated by the user, the system must "pre-apologize" by making transparent its thought process to some extent. At the same time, it must appear reliable by showing confidence in what it has suggested, and of course by minimizing mistakes and user effort. These factors are necessary in order for the user to depend on the system and trust it, because without that trust the entire scheme is for naught. This attainment of trust remains the ultimate relational goal of the system, encapsulating both its requirements and needs for success.

Early User Feedback

Unfortunately, I was unable to schedule time with the C*D office members until now. However, I have had the chance to talk to some sponsors (potential, established, and new). Additionally, focus groups with those involved in the system design and development has also provided a number of useful insights.

The short lead-time fairly precluded longer sessions with the sponsors. However, I was able to speak to representatives from Cisco (new sponsor), Accenture (prospective, or some would say "ever-interested", sponsor), and Sony (old sponsor, but delegation was composed of first-time visitors to the lab). Firstly, the level of confusion about what part of the lab was engaged in what specific activity appears to be universal. Additionally, the sample highlights the fact that even long-time sponsors will likely have many new members here for the event. The first two visitors were somewhat uncertain of the capability of the system to direct them well, which may indicate a prevalent mindset among those more familiar with the lab. However, at the same time they believed that if the system were of use regardless of the accuracy of the system, it would garner attention regardless. In the context of the lab, all visitors rightly envisioned the system to be something like the C*D personnel who escort them between locations. While this is a reasonable parallel, the essential issue is that the system has a reasonable level of knowledge about the actual work, and works in a proactive manner. I believe that this distinction may not become clear to users until the system is put into effect. The visitors from Sony expressed a sense of novelty, which will also likely be prevalent among the many newcomers at the TTT event.

This will effectively work in favor of the system, because the novelty factor will compel users, at least initially, to interact with the system.

Several focus meetings have also come to pass which have further defined goals of the project. One of the primary problems has been the information about the Media Lab that is user-visible in phase two, and underlies the entire project. Extensive analysis of the PLDB entries (using common sense reasoning tools as well as low level tools for structural analysis and gisting) has resulted from the conclusions of some of the initial meetings, indicates that there are three problems that are pervasive. Firstly, hardly any of the descriptions talk about what the projects are. Instead, the concentrate on the things the projects do. While this is not a major problem for members of the lab, since we have a large amount of context, it is a great hurdle to outsiders. Moreover, when stripped of linguistic and semantic context into keywords, the words that are considered "key" have hardly any meaning remaining in them at all. This is also clear a problem present in the keyword lists given to the sponsors. The second problem is that the work that we do in the Media Lab is highly specialized. This causes a high level of nuances inhomogeneity even within groups. In essence, the PLDB entries try to highlight these nuances. This has a dual repercussion of preventing the use of existing keyword ontologies (such as the one used by ACM) for classification, and causing the keywords to be inhomogeneous enough so that even with manual selection of keywords, associativity is too minimal for use in preference inference. Lastly, the PLDB as it stands today is generated by a computer, and structurally, it is easy to read for a computer, and almost impossible to read for a human being. A raw dump of the database entries would be approximately as simple for people to parse. Stripped of the complex interconnections of the Media Lab "no walls" policy, the entries themselves lack the associative context needed for humans to understand the whole of the lab. After several prototype test and discussions (all done last week, leading to the aforementioned time crunch), a new tactic has been adopted which involves less automated parsing, but at the same time prevents the proliferation of the above problem. Once finalized, the system will be made available to media lab members for data entry.

Keeping in mind the suggestions from the interview, as well as focus meetings on the applications that will be running on the badge, several other functionalities have been migrated from the badge to the system. The system will now allow users to find each other at the kiosks instead of using the badge based searching method. Additionally, it will perform real time clustering of the users based on their actions, and assign them unique icons that would form impromptu cliques with similar interests. This is a compromise from actually outputting a shared interest, which would tend to distract users from the actual task-at-hand of interacting with others. This provides the system with two new uses, with one more frivolous and low cost than the other to provide a gradient in terms of relationship management. These added functions increase interactions with the system, thus also increasing probability that the user with then use the other available features. The search feature requires that the system track users fairly accurately, which in turn allows an increase in search relevance by

de-prioritizing locations the user has already visited or visited recently. Additionally, it is no longer required that the user vote in order to check for compliance, which should provide a good tool for evaluating success of the system.

Initial Spec

I initially described system as having three phases. These phases concern the deployment schedule of the system. I have covered the responses to the initial interviews and focus groups above. However, I would like to reiterate the design goals of the three deployment phases, which will essentially consist of completely separate systems connected only by the underlying data.

The first phase will be a website with minimal relational qualities beyond the level of "people engineering" through design. It will be designed towards the goals stated in the first section while maintaining the highest possible levels of usability. The goal here is to have the Media Lab community create a high quality, well associated map of the work being done.

The second phase will be similar to the first phase, except that the focus will be somewhat different, as stated previously. The core difference will be the input modality and "personality" projected by the system, which will focus primarily on keeping the user from stopping use of the system until customization is complete. Almost all other goals will be secondary to this, because a failure at this phase will undermine the next, and it is the one most susceptible to noise from external pressures.

The relational aspects of the phase three recommender system will be the most pronounced. The system will be modeled via a natural language agent that will translate the inference patterns of the core recommender into a simple and concise explanation of the actions of the system. It will attempt to be as conversational as possible. At the same time, it will attempt to convey to the user its ability to learn and adapt, which should allow the user to place greater trust in the system to improve than one would with a rote system.

Demonstration

This project will be living by the "demo or die" motto, and will be under full lab-wide deployment starting mid-April for phase one. All parts must be (and thus will be) fully implemented for the demo to work. The first two parts will be implemented using PHP and accessed via the web. The last section will be implemented in Java. The client binaries will be used to set up kiosks throughout the lab. Messages about the kiosk setup will be sent to advisors shortly, along with information for completing input for phase one.

Evaluation

The system is now able to track user locations with moderate accuracy (room-level). The recommendations and location data will be stored for later analysis, and compliance data will be extracted and correlated with voting data from the badges. The sponsors will afterwards have access to a website containing information about projects and persons they were interested in, and will be presented with a survey about their relationship with the system and level of trust garnered (primary metric of interest in this case). This will reduce the problems encountered by surveys handed out at sign-out, and will have the added bonus of providing an incentive for sponsors to complete the information (ie, it will be on the way to the information they requested). As usual, they will have the option to decline or delay answering the questionnaire. These three metrics (compliance, voting, and questionnaire) should together provide a reasonable understanding of the shortcomings and successes of the system.