Computing point-of-view

By Hugo Liu

Thesis Proposal for the degree of Doctor of Philosophy
at the Massachusetts Institute of Technology

December 2005

Professor Pattie Maes
Associate Professor of Media Arts and Sciences
Massachusetts Institute of Technology

Professor William J. Mitchell
Head, Program in Media Arts and Sciences
Alexander W. Dreyfoos, Jr. (1954) Professor
Professor of Architecture and Media Arts and Sciences
Massachusetts Institute of Technology

Professor Warren Sack
Assistant Professor of Film & Digital Media
University of California, Santa Cruz
Computing point-of-view: modeling and simulating judgments of taste
Hugo Liu
Media Arts and Sciences, MIT
hugo@media.mit.edu

December 2005

Abstract
Point-of-view affords individuals the ability to judge and react broadly to people, things, and everyday happenstance; yet it seems ineffable and quite slippery to articulate through words. Drawing from semiotic theories of taste and communication, this proposal presents a computational theory for representing, acquiring, and tinkering with point-of-view. I define viewpoint as an individual’s psychological locations within latent semantic “spaces” that represent the realms of taste, aesthetics, and opinions. The topologies of these spaces are acquired through computational ethnography of online cultural corpora, and an individual’s locations within these spaces is automatically inferred through psychoanalytic readings of egocentric texts. Once acquired, viewpoint models are brought to life through viewpoint artifacts, which allow the exploration of someone else’s perspective through interactivity and play. The proposal will illustrate the theory by discussing interactive-viewpoint-artifacts built for five viewpoint realms—cultural taste, aesthetics, opinions, tastebuds, and sense-of-humor. I describe core enabling technologies such as culture mining, common sense reasoning and textual affect sensing, and propose a framework to evaluate the accuracy of inferred viewpoint models and the affordances of viewpoint artifacts to recommendation, self-reflection, and constructionist learning.

1 Introduction
Our capacity for aesthetics and emotional reaction is one of the most celebrated bastions of humanity. Underlying our explicit knowledge and rationality is a faculty for judgment—the capacity to prefer, to view the world through our individual lenses of taste. An interesting intellectual question is: can a computer model a person’s taste, aesthetics, and opinions deeply enough to predict their judgment? The proposed thesis explores this question, in depth.

The user modeling literature has been exploring predictive models of persons for over two decades. Two main approaches have emerged—stereotype-based models and behavior-based models. Stereotype-based models, such as Elaine Rich’s (1979) book recommender system, represent persons by demographic categories and acquire each user’s profile by asking a set of questions. Behavior-based approaches perform statistical inference over a history of a user’s actions to predict future actions—seemingly examples are collaborative filtering (Shardanand & Maes 1995), and
Bayesian goal inference (Horvitz et al. 1998). While stereotypes are potentially deep models, they must be handcrafted through exercises of superb intuition, and even then, stereotypes tend to under-fit the individuality of people. Behavioral models can be acquired automatically, but the models themselves tend to be shallower, and there is a danger of over-fitting. In the literature up to now, behavior-based approaches have almost exclusively been used to model individuals within the context of using specific applications, like tutoring systems and shopping websites. As such, these ‘user’ models do not capture persons with much generality. While user modelers routinely capture user’s ratings of items within application contexts, a general model and simulation of a person’s tastes, aesthetics, and opinions that cuts across application domains has not yet been achieved.

Scope. The scope of the proposed thesis, then, will be to approach the computational modeling of a person’s taste, aesthetics, and opinions in more sophisticated ways, modeling persons as wholes, not just as application users. I pose things like taste, aesthetics, and opinions, as points-of-view in order to emphasize this crucial metaphor—we are always judging the world through the optics of some viewpoint, and our viewpoint can be seen as a location within the greater cultural space of possible viewpoints. I also differentiate the treatment of point-of-view in this thesis from previous computations of point-of-view by Warren Sack (1994; 2001). Whereas Sack’s robotic readers mine ideological ‘spin’ structures from news stories, this thesis examines psychological point-of-view. Ideological point-of-view is a set of politicized and institutional conventions, what Lakoff (Lakoff & Johnson 1980) calls metaphorical framings, e.g. the Islamic martyrs versus the Islamic terrorists; psychological point-of-view is concerned with modeling the interior experience of one individual—how a person sees the world idiosyncratically by possessing various unconscious, culturally-conditioned lenses that give emotional tint to judgments and reactions.

Claims. The thesis will present a series of experimental systems that have been built to capture and simulate psychological viewpoint under five realms—cultural taste, aesthetics, opinions, tastebuds, and sense-of-humor. These experiments will support the thesis’ three main claims, listed below.

- **Viewpoint can be modeled and simulated as an individual’s psychological locations within latent semantic spaces that represent cultural taste, aesthetics, and opinions.** This claim will be dissertated as a computational theory of point-of-view, building closely on existing Semiotic theories of viewpoint, aesthetics, culture, and taste (Jung 1921; Barthes 1964; Bourdieu 1984; Latour 2005).

- **The topology of viewpoint spaces can be acquired by semantic mining of large-scale cultural corpora; while an individual’s location can be inferred by psychoanalytic reading of his/her egocentric texts.** Viewpoint spaces are acquired by a technique called culture mining, which combines natural language processing and machine learning to discover the emergent semantics of
cultural corpora like social network profiles and weblog communities. In Semiotics, psychoanalytic reading means reading text deeply in order to model the author (Silverman 1983). I extend existing computational models of reading (Zwaan & Radvansky 1998; Moorman & Ram 1994) to acquire the author’s viewpoint location. Common sense reasoning, and textual affect sensing are two technologies critical to computing psychoanalytic reading.

Interactive viewpoint artifacts that simulate a person’s taste judgments can provide deeper user models for recommendation, and can support constructivist learning of other people’s perspectives. Interactive artifacts have been implemented and evaluated for each of the five experimental systems. The artifacts are interface agents that perform just-in-time information retrieval (Rhodes & Maes 2000)—in other words, they observe the user’s context of browsing and writing, and constantly offer their reactions to the user.

A process model (Figure 1) contextualizes the three claims within what is to be achieved by this sort of modeling. For the past four years, I have been implementing and evaluating these five experimental systems and the core natural language and common sense components. More recently, I have reflected on the interconnections between these various systems and I now propose

**Figure 1.** A process model of the proposed technique for viewpoint modelling.
to organize and unify their discussion under the banner of point-of-view models. Deep person modeling fits well under the banner of Ambient Intelligence Group, where our motto is insight, inspiration, and interpersonal communication. Other than completing several further evaluations, I see my main research task as synthesizing together a coherent theoretical framework that will make contributions to the User Modeling and Artificial Intelligence literatures. The problems of viewpoint, taste, and aesthetics are so complex that to explain them would necessarily require crossing many literatures. Thankfully, Semiotics is a virtual literature whose mandate is the linguistic modeling of self, culture systems, and aesthetics. Semiotics is a stew of other disciplines such as literary theory, psychology, psychoanalysis, sociology, and cognitive science. Thusly, the major aspiration for this thesis is to import many important deep Semiotic theories of viewpoint, taste, and aesthetics into the computational literature; and to substantiate, through extensive evaluation, the efficacy of these theories in accomplishing sophisticated modeling of persons.

2 Proposed Research
Section 2.1 presents a computational framework for point-of-view. Section 2.2 discusses three core technologies necessary for viewpoint computation—culture mining, common sense reasoning, and textual affect sensing. Section 2.3 overviews the five implemented experimental systems and their viewpoint artifacts, which are already implemented. Section 2.4 outlines an evaluation strategy for this thesis.

2.1 Computational Framework

Groundings. I compute viewpoint as an individual’s psychological location within latent semantic spaces such as cultural taste, aesthetics, opinions, tastebuds, and humor1 (Figure 2a). This framework is grounded in the Semiotics literature’s tradition of psychological situationalism (Hume 1748) and social constructionism (Lacan 1957; Bourdieu 1984; Latour 2005)—the notions that individuals are constructed by their environment, and that subjectivity is the product of socialization. Pierre Bourdieu’s Distinction: A Social Critique of the Judgment of Taste (1984) is a seminal work in Semiotics which needs to be mentioned upfront, for it comes very near to being a direct theoretical basis for the computational framework presented in this thesis. In that work, Bourdieu surveyed 1200 French persons in the 1960s, computed statistical correlation, and found a relationship between taste and class structure in French society. He theorized an individual’s judgment faculty as being structured by a set of personal dispositions called a habitus, which is constituted from a cultural field of socio-economic conditions. The intersection of the personal habitus and cultural field is called doxa—doxa, then, is

---

1 I do not mean to indicate that these five viewpoint realms offer a complete or canonical account of an individual’s point-of-view. Rather, they are examples of some viewpoint realms which might be considered by common sense to be significant.
the site of the individual’s cultural identity. Habitus, field, and doxa, I suggest, is almost a parallel vocabulary for viewpoint, space, and location, respectively. Space/field defines the limits of what is possible. Location/doxa defines where an individual’s psychology fits into the culture. Viewpoint/habitus is an individual’s system of dispositions (e.g. system of opinions, system of aesthetics, system of taste); this is the psychological structure that can be used directly to predict the individual’s future judgments and reactions.

Building on the success of Bourdieu’s theory, the computational framework presented here considers more than just the space of cultural taste (Taste Fabric / InterestMap) —it extends the space/location/viewpoint metaphor to experiment with modeling persons under other realms such as perceptual aesthetics (Aesthetiscope), opinions (What Would They Think?), tastebuds (Synesthetic Recipes), and humor (Buffolo). The topology of these spaces can often be acquired through computational ethnography of online cultural corpora—the invocation of latent semantic mining to reveal the emergent correlations and network structure of a cultural space. For example, Taste Fabrics is a densely connected network of cultural taste, mined from automated analysis of the texts of 100,000 social network profiles.

An individual’s locations within these spaces can often be inferred through psychoanalytic readings of egocentric texts (self-revealing, self-describing), for example, a diary, a research paper, a social network profile. Psychoanalytic reading means reading not for the message, but for the subjectivity of the message sender. The technique of psychoanalytic reading is anchored in Semiotics—Roman Jakobson’s theory of communicative function (1960), JL Austin’s speech acts theory (1962), and Kaja Silverman’s suture technique for psychoanalyzing narratives (1983). The common ground of these theories is that they all pose emotional attitude as the unifying force of subjectivity. In speech acts, underlying each utterance is the illocutionary force, which is the author’s emotional posture, such as aggression, agreeableness, or sadness. Similarly, Jakobson suggest that the goal of emotive communication is to paint a portrait of the author, which is why the present research prefers emotionally expressive egocentric texts as a way to ensure that the subject can be modeled. Heeding these theories, the proposed thesis computes psychoanalytic readings by reading for the unconscious emotional undertone of topics discussed in egocentric text. Natural language understanding, common sense reasoning and textual affect sensing are core technologies which achieve psychoanalytic reading.

**Knowledge representation for viewpoint spaces.** Figs. 1b-1d illustrate three varieties of knowledge representation used in this thesis research to model latent semantic spaces. But why different representations of viewpoint and not one? Because sometimes the space has straightforward dimensionality (Fig. 2b) while other times a space can appear quite disorganized (Figs. 2c-d). The choice of representation is ultimately an engineering consideration, but I believe that the three representations developed through this thesis are principal.
There is a pecking order. Dimensional spaces are most preferred, as meaning is most organized, and Cartesian distance is easily measured. The viewpoint space for perceptual aesthetics (Fig. 2b) developed in this thesis is dimensional—its axes are based on Carl Jung’s theory of fundamental psychological functions (1921). Next best are semantic fabrics, which are \( n \times n \) correlation matrices with topological features like cliques and stars. Semantic fabrics are fully connected representations, but with only patchwork consistency—distance is non-Cartesian here but can be measured simply by spreading activation (Collins & Loftus 1975). The mining of the latent space of cultural tastes from social network profiles (Liu & Maes 2005a, Liu, Maes & Davenport 2006) leverages semantic fabrics because while the mutual information between cultural products (e.g. books, music, films) can be calculated, it is believed that the
The dimensionality of this space are too complex to be able to name principle dimensions. Still, the space enjoys partial organization such as cliques of highly correlated products, and star structures around “identity hubs” (e.g. products like ‘yoga’, ‘hiking’ can be organized around the hub of ‘new agers’). In the poorest case, neither dimensions nor connectedness are known, as is the situation for this thesis’s modeling of a person’s system of opinions. The space of all possible opinions (opinion = an attitude about a topic) is consistent around a few ideological centers like politics and academia, but there is no obvious global consistency. My What Would They Think? system (Liu & Maes 2004) develops a semantic sheet representation (Fig. 2d) — to make the best of this situation.

Inspired by Marvin Minsky’s “causal diversity matrix” (Minsky 1992), Figure 3 summarizes these representational tradeoffs. Note that a third dimension could also be name — semioticity. We could distinguish “dimensional spaces” as being either a semiotic /structuralist space like Jung’s modes of perception, or as being a data-emergent “quality space” (Gärdenfors & Holmqvist 1994).

**Organizing Principles of Viewpoint.** Consistency gives shape to viewpoint space. Without consistencies, applying viewpoint models to predict reactions to arbitrary stimulus, or *fodder*, would just have to resort to rote dictionary lookups. If the answer is not in the dictionary, no reaction could be given. Dimensional spaces like Jung’s perceptual dimensions have built in consistency. However, loci of consistency in Semantic Sheets and Semantic Fabrics need to be found opportunistically. Identifying promising loci of consistency for various viewpoint realms is a contribution of this thesis. For cultural taste space, I nominate three topological organizing features — identity hub-and-spokes, taste-cliques, and taste neighborhoods,

**Figure 3.** Semantic diversity matrix. Point-of-view spaces can be conceived in terms of their consistency and connectedness — for each case, an appropriate knowledge representation is specified. The top row is semiotic-symbolic in quality; the bottom row is ethnographic/connectionist in quality.
discussed elsewhere (Liu, Maes & Davenport 2006). For opinion space, four forces of consistency are nominated—1) Minsky’s *imprimer theory* (Minsky, forthcoming) predicts that a person’s opinions are partially structured by opinions of their parents and mentors; 2) ideology of politics, and academia structure opinions such that if a person has a positive opinion toward Social Security, that has many ideological entailments; 3) folksonomies of topics imply underlying consistency along topic-subtopics inheritance trees, e.g. attitude toward “macramé” predicts attitude toward “crafts”; and 4) analogical reasoning (Gentner 1983; Fauconnier & Turner 2002) can be used to predict reactions to unknown fodder by structure-mapping to identify similar things, e.g. attitude toward “rocks” can be predicted by attitude toward “trees” by shared conceptual intensions (*sic*). Finally, techniques from *truth maintenance systems* (Doyle 1980) are applied to maintain patchwork consistency, though contradictions do occur and these are presented as “soft-constraints.”

**Simulating Viewpoints.** It may be said that just as light has no resting mass, point-of-view is not intelligible in stasis. To fully appreciate and understand a viewpoint, its space+location must be animated and allowed to react to a broad many things.

Simulating judgment means applying the location in space data, to create a reaction to some arbitrary semantic stimulus called fodder. Analogy (Gentner 1983; Fauconnier & Turner 2002) and context-biased spreading activation (Collins & Loftus 1975; Liu 2003) are chief techniques to achieve this reasoning. Although with viewpoint models we go beyond rote memory-based application of old ideas to new fodder, viewpoint simulation is still not capable of applying viewpoint models in any particularly *clever* way to new situations. Humans are capable of evolving their viewpoint nimbly as new fodder presents opportunities for belief revision, but machines are not yet capable of simulating the complex *dialectic process* (Bakhtin 1935), which may affect judgment. A goal for the thesis is to discuss how the simulation of viewpoint could become dialectical, how an artificial viewpoint could contradict and overcome itself cleverly—what Hegel calls *Aufhebung* (1807). Viewpoint models and simulation carry specific implications for dialectics—a central problem in critical theory. If Aufhebung could be simulated, it would represent a major breakthrough for the computation of inspiration.

To animate computed viewpoint models, viewpoint artifacts are created—such as the Identity Mirror (Liu, Maes & Davenport 2006; Liu & Davenport 2005), the Aestheticscope (Liu & Maes 2005b; Liu & Maes 2006), virtual mentors in What Would They Think? (Liu & Maes 2004), and avatars in Synesthetic Recipes (Liu, Hockenberry & Selker 2005). Viewpoint artifacts reify space+location models by having them constantly react *just-in-time* and *just-in-context* to a broad range of fodder put forth to them implicitly or explicitly by a user, and by visualizing these reactions through visual metaphors. Furthermore, each viewpoint artifacts allows for tinkering, play, and explanation, e.g. virtual mentors can “justify” their reactions with
quotes, and identity can be negotiated in the Identity Mirror by a “dancing” interaction in front of the mirror. The importance of tinkering is likely due to the fact that a reaction’s motivation cannot be easily grasped without exploring the immediate context and conditions surrounding the reaction.

2.2 Core Enabling Technologies
Three core technologies that drive the acquisition of viewpoint models from machine readings of text are culture mining, common sense reasoning, and textual affect sensing. Machine learning techniques and hand engineering of many support semantic knowledge bases are also important, but they are not discussed here.

Culture mining. In Roland Barthes’ (1964) semiotic model of culture, he proposed cultures to be the set of symbols salient to the unconscious of a population. He said also that these symbols are organized into semantic systems and have valence, or degrees of privilege. Similarly, Clifford Geertz (1973) remarked that cultures were ‘webs of significance’ which implicated people into them. From Barthes and Geertz’s representation of culture, two of the most definitive ever presented, I define the culture mining problem as uncovering the symbols, interconnectedness, and significance from a cultural corpora, such as a corpus of social network profiles, or a corpus of conservative versus liberal news texts. The technique for culture mining is *computational ethnography*—a combination of automated language analysis to extract significant symbols, and machine learning to statistically infer the latent dimensionality and connectedness of symbols.

Relevant machine learning and statistical language modeling techniques include—Latent Semantic Analysis (Deerwester et al. 1990), Support Vector Machines (Joachims 1998), Multi-Dimensional Scaling (Kruskal & Wish 1978), and Principle Components Analysis. Relevant language analysis tools solve problems present in the unstructured natural language nature of many online cultural corpora—including discourse segmentation, tokenization, named-entity recognition, spelling correction, part-of-speech tagging, deixis resolution, phrase chunking, linking, gisting syntactic, semantic, and thematic role frames, natural language generation, topic spotting, summarization, and statistical language modeling. For the bulk of these language tasks, I have developed a natural language understanding platform for Python, called MontyLingua (Liu 2002)—now widely used since my releasing it to the Computational Linguistics and AI communities.

Commonsense reasoning. Commonsense reasoning is a core component of machine readers that will read texts to acquire viewpoint spaces and locations. The essential insight that distinguishes machine reading—or Story Understanding / Narrative Comprehension as it is also called—from mere deep text parsing is that more than what a text explicates, it also implies and insinuates through *subtext*, and it requires contingent knowledge in the form of *backtexts* to decipher the full meaning of an utterance. To read
subtexts and with backtexts, the Artificial Intelligence community has applied approaches such as Schankian scripts and plans (Schank & Abelson 1977), and more recently, large scale databases of world knowledge (Lenat 1995; Mueller 2000; Singh et al. 2002). The proposed thesis uses the latter approach as it gives broader semantic coverage—a feature necessary to the interpretation of domain-independent texts.

Cyc (Lenat 1995), ThoughtTreasure (Mueller 2000), and Open Mind Common Sense (Singh et al. 2002) are three approaches to large-scale common sense knowledge acquisition and reasoning. Cyc and ThoughtTreasure have logical representations and are more suitable for rigorous deep reasoning about situations, while Open Mind Common Sense and its ConceptNet (Liu & Singh 2004b) has a natural language representation, and thus excels at contextual reasoning over natural language texts (Liu & Singh 2004a). ConceptNet is a semantic network of common sense facts, with built-in methods for contextual expansion and analogy.

Examples of use in this thesis are as follows. The conceptual analogy faculty of ConceptNet is used to apply viewpoint models to predict reactions to unknown concepts in WWTT (Liu & Maes 2004) by situating the unknown fodder into the space of known concepts, also called conceptual alignment in the Cognitive Science literature (Goldstone & Rogosky, 2002). In the aesthetic viewpoint space, ConceptNet’s getContext() feature is used to brainstorm the rational entailments of a text, in order to generate the “shadows” that a fodder casts onto the “Think” axis. Finally, ConceptNet is a principle component of another core technology—textual affect sensing.

**Textual affect sensing.** Judgment is the behavioral and measurable expression of viewpoint, and the primary quality of judgment is affect. In fact, Ortony, Clore and Collins (1988) concisiated the definition of “emotion” to mean the expression of an affect about a person, thing, or event. Emotion and judgment thus can be represented basically as the bound pair (thing, affect). In some of the viewpoint systems to be presented in this thesis, affect manifests as choice implicature. For example, in the cultural identity space acquired through linguistic ethnography over social network profiles, individual choose to display certain items into their profile of “my favorite things,” and that choice can be viewed as a judgment act (Austin 1962; Habermas 1981) which says that things listed in the profile are more pleasurable and arousing and dominated over than things not listed in the profile.

Other times though, affect must be inferred from unstructured natural language texts—for example, the machine should learn from the utterance “my mother is a loving and generous woman” that the speaker judges his mother positively. To complete this task, a topic spotter looks for the topics present in sentences, paragraphs, and documents, while a textual affect sensor appraises the affective qualities of each segment of text. Binding those two outputs to each other as (topic, affect) pairs, and using classical reinforcement learning (Kaelbling, Littman & Moore 1996) to generalize stable
(topic, affect) pairs from training data, we have the beginnings of a model of a person’s system of attitudes/opinions.

To accomplish comprehensive textual affect sensing, I sense separately surface and deep affect. Surface, or rhetorical affect, can be measured as word-choice; I sense it by combining the Sentiment headwords of Roget’s Thesaurus (1911), a corpus of psychologically normalized affect words called ANEW (Bradley & Lang 1999), and an affective lexical inventory produced by Ortony, Clore and Foss (1987).

Deep affect is the pathos permeating from the contingent imagined consequences of an utterance and can be communicated without mood keywords at the surface. For example, the utterance “I was fired, my wife left me, and she took the kids and the house” uses no surface keywords to nonetheless convey a negative affect quite powerfully. Deep affect sensing is attempted using Emotus Ponens (Liu, Lieberman & Selker 2003), a textual affect sensor built using the Open Mind Common Sense corpus (Singh et al. 2002). The basic idea is when the affect of a concept is unknown, it can be approximated by the affect in its surrounding conceptual neighborhood. For example, supposing that the concept “get fired” is not annotated with affect, ConceptNet (Liu & Singh 2004b) has semantic links which connects “get fired” to other nodes which are annotated with affect such as “recession” (probable cause), “stupid person” (probable cause), “no money” (probable consequence), “hungry” (probable consequence). Thus the affect of “get fired” can be guessed by its context.
Table 1: A summary of the implemented experimental systems.

<table>
<thead>
<tr>
<th>experimental system</th>
<th>cultural tastes</th>
<th>perceptual aesthetics</th>
<th>opinions</th>
<th>tastebuds</th>
<th>humour</th>
</tr>
</thead>
<tbody>
<tr>
<td>TasteFabrics/InterestMap</td>
<td>Aestheticscope</td>
<td>WhatWouldTheyThink?</td>
<td>Synesthetic Recipes</td>
<td>Buffalo</td>
<td></td>
</tr>
<tr>
<td>Bourdieu’s Distinction</td>
<td>Jung, Dewey</td>
<td>Ortony, Barthes, Minsky</td>
<td>Synesthesia, Davidson</td>
<td>Freud’s Hydraulics</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>space representation</th>
<th>Semantic Fabric</th>
<th>Dimensional Space</th>
<th>Semantic Sheet</th>
<th>Semantic Sheet</th>
</tr>
</thead>
<tbody>
<tr>
<td>100,000 social network profiles</td>
<td>Jung’s functions +</td>
<td>e.g. left v. right</td>
<td>food folksonomies,</td>
<td>corpus of ethnic</td>
</tr>
<tr>
<td>space modelled from</td>
<td>AestheticReader s</td>
<td>~wing newspaper corpora</td>
<td>60,000 recipes, OMCS</td>
<td>jokes</td>
</tr>
<tr>
<td>location modelled from</td>
<td>psychoanalytic reading</td>
<td>psychoanalytic reading</td>
<td>psychoanalytic reading</td>
<td>psychoanalytic reading</td>
</tr>
<tr>
<td>a person’s social network profile</td>
<td>for MBTI, in ESCADA</td>
<td>for (affect, topic)</td>
<td>for (affect, topic)</td>
<td>reading for (pressure,</td>
</tr>
<tr>
<td>their ethos in the fabric</td>
<td>(L.u &amp; Mueller</td>
<td>pairs in egocentric</td>
<td>pairs in egocentric</td>
<td>topic) pairs in</td>
</tr>
<tr>
<td>(forthcoming)</td>
<td>texts</td>
<td>texts</td>
<td>texts</td>
<td>weblogs.</td>
</tr>
<tr>
<td>kind of reasoning used to simulate</td>
<td>Pandemonium,</td>
<td>Reflexive memory,</td>
<td>Spreading activation</td>
<td>Reflexive memory,</td>
</tr>
<tr>
<td>judgment</td>
<td>AestheticReaders and</td>
<td>analogy, Minskian</td>
<td></td>
<td>analogy</td>
</tr>
<tr>
<td>(sic)</td>
<td>color logics</td>
<td>imprimers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>interactive artifact</td>
<td>IdentityMirror lets you</td>
<td>Aestheticscope uses</td>
<td>A panel of virtual</td>
<td>A joke tells uses</td>
</tr>
<tr>
<td>see your reflection in cultural space</td>
<td>see your reflection in</td>
<td>your aesthetic</td>
<td>mentors playing back</td>
<td>your Freudian</td>
</tr>
<tr>
<td>of taste</td>
<td>cultural space of taste</td>
<td>viewpoint to create</td>
<td>their opinions to</td>
<td>humour model to</td>
</tr>
<tr>
<td></td>
<td></td>
<td>art customized to words</td>
<td>you just -</td>
<td>select jokes for you</td>
</tr>
<tr>
<td></td>
<td></td>
<td>/poetry / lyrics</td>
<td>in-time while you</td>
<td>based on your</td>
</tr>
<tr>
<td>task-based evaluations</td>
<td>*viewpoint-driven</td>
<td>*efficacy of viewpoint</td>
<td>*using WWTT to learn</td>
<td>culture + mood</td>
</tr>
<tr>
<td>(*indicates already completed)</td>
<td>recommendation</td>
<td>driven art creation</td>
<td>rapidly about a mentor</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>utility in helping</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>mum/dad decide what</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>to make for her</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>family</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>efficacy of viewpoint-</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>driven joke</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>recommendation</td>
<td></td>
</tr>
</tbody>
</table>
2.3 Already Implemented
This section describes the already implemented experimental viewpoint systems and their associated interactive viewpoint artifacts (Table 1). These will illustrate the computational framework and theoretical principles already enounced. The artifacts introduced here have implications for diverse applications such as technological support for self-reflection, perspectival tools for learning from others, interfaces for visualizing and searching human narrative content, psychographic visualizations for marketing and ethnography, and so on.

2.3.1 Major Examples

(Opinion Space) What Would They Think? (Liu & Maes 2004) is a system for modeling personal attitudes and the space of opinions at large using the Semantic Sheet representation shown in Fig. 2. A user can build a new “persona” by supplying an icon and pointing the system to some egocentric texts that are self-revealing and self-describing—i.e. position papers, instant messaging logs, emails, weblogs. The system reads and infers from the text a system of attitudes for that persona. Personae are embodied into virtual mentors (Fig. 4a) who continually observe the user’s browsing and writing activities, offering up just-in-time and just-in-context feedback to the user’s “fodder” through visual metaphors. To find out why a mentor reacted in a particular way, mentors can be double-clicked to pop up an explanation window—this window displays a list of quotes snipped from the mentor’s “memory” of egocentric texts, rank-ordered by how well they justify the reaction that was given. For example, virtual mentor Roz Picard reacts

![Figures 4a-b. What Would They Think?](image)

What Would They Think? is a panel of virtual mentors who continually observe the user’s browsing and writing activities, offering up just-in-time and just-in-context feedback to the user’s “fodder”. Visual metaphors: red=> displeasure, green=>pleasure, dim=>unaroused, lit=>aroused, sharp=>dominant, blurry=>submissive. 3a) (left) depicts a panel of AI luminaries reacting to the user’s surfing of the Social Machines Group website. 3b) (right) shows a Democratic Party persona and a Republican Party persona (trained on their party talking points) reacting to an article entitled, “What’s Wrong with the Contract with America?”
negatively to the utterance “Robots will have consciousness” which is defended with quotes like “Several of my colleagues believe it’s just a matter of time and computational power before machines will attain consciousness, but I see no science nuggets which support such a belief.” Fig. 4b depicts the modeling of two cultures qua personae. In WWTT, cultures can be treated commensurately with individuals. The proposed thesis will pre-generate a fabric of cultural opinions to acquire the opinion space. Using this opinion fabric, individuals can be located as inhabitants of particular cultural opinions by applying simple alignment or “diff” techniques between cultures’ reactions and individuals’ reactions.

(Perceptual Aesthetic Space) The Aesthetiscope (Liu & Maes 2005b) is an art robot that renders color grid artwork a la Ellsworth Kelly and early Twentieth Century abstract impressionists (Figure 5). A model of the user’s perceptual aesthetics guides the manner and quality of the generated artwork. The perceptual aesthetic space (shown in Figure 2b) is modeled as having the five dimensions of Think, Sense, Intuit, Feel, and Culturalize—these dimensions are based on Carl Jung’s fundamental modes of perception (1921). Though not yet implemented, the proposed thesis will automatically acquire the user’s aesthetic viewpoint through readings of egocentric text. Currently these dimensions must be specified manually. As a perspectival artifact, the Aesthetiscope reacts to “fodder” given to it.

![Figure 5. Aesthetic viewpoint driven rendition in the Aesthetiscope. The left column shows how the art robot renders the aesthetic impression of the words “sunset” (above) and “war” (below) through the viewpoint of a Realist (e.g. Sense=90%, Think=60%, Culturalize=40%, Feel=20%, Intuit=10%). The right column shows the same fodder rendered through the viewpoint of a Romantic (e.g. Sense=50%, Think=20%, Culturalize=70%, Feel=90%, Intuit=80%).]
such as a word, a poem, or song lyrics. For example, it continuously observes what poetry the user is reading or what songs are queued in the playlist, dynamically changing the color grid artwork to “pair” with the fodder, just as wines are selected to pair with a cheese course. Another perspectival game that can be played is for two individuals both standing in front of the same artwork visualizing some poem to find their shared aesthetic (by averaging their locations), or to violated each other’s aesthetic (by allow one aesthetic viewpoint to corrupt another viewpoint). I am particularly interested on how deeply held aspects such as aesthetics can be exhibited or worn on one’s sleeve so to speak, like a piece of clothing avails identity and taste.

(Cultural Identity & Taste Space) Identity Mirror (Liu, Maes & Davenport 2005; Liu & Davenport 2005) is a mirror to support self-reflection that lets you “see who you are, not what you look like.” As shown in Fig. 6, the mirror’s computed reflection overlays a swarm of keyword descriptors over an abstracted image of the “performer.” The performer can use dance to negotiate his identity—for example, walking to and fro the mirror affects the granularity of the keywords being shown, which describe a far away

![Figure 6. Self-reflexive performance with the identity mirror. A swarm of keywords shows a user’s situation within the cultural fabric of identity/taste, and with respect to the attentional biases of the zeitgeist as calculated by monitoring daily news streams. The user’s social network profile is used to locate the user within the cultural fabric.](image-url)
performer using broad strokes like subculture keywords (e.g. fashionista, raver, intellectual, dog lover), but describe an up-close performer with descriptors like song names, books, food dishes, etc. When movement is slow and deliberate, the keywords more semantically distant from the performer’s ethos appear in the computed reflection, but those keywords are quickly dashed with sudden movements.

The Identity Mirror uses a social network profile to locate the performer’s viewpoint within the cultural fabric of identity and taste.

The cultural “taste fabric” (Liu, Maes & Davenport 2005) is derived by computing the latent semantic connectedness of “interest keywords” (music, books, sports, subcultures, etc) from analysis of the texts of 100,000 social network profiles. The performer’s location on the fabric is calculated by reading his social network profile, mapping that profile onto the nodes of the fabric, and using spreading activation (Collins & Loftus 1975) to define an ethos (a weighted collection of nodal activations). In the mirror artifact, the identity/taste viewpoint of the performer is visualized as a swarm of keywords. The viewpoint “reacts” to changes in the daily news stream. For example, around the time of the summer Olympics, the sports-centered news wire would bias the cultural fabric by highlighting nodes relating to the Olympic sports. The reflection in the mirror simulates the performer’s viewpoint by selectively interpreting the new cultural situation, and displaying just what exists at the intersection of the performer’s ethos and the news-du-jour’s ethos. Ambient Semantics, another system using the taste fabric, is an artifact that uses viewpoint to predict whether or not one individual would find another person to be sympathetic.

2.3.2 Minor Examples
These experimental systems will not be used to introduce the computational framework. Rather they will be used to thicken descriptions of the framework, to raise questions, and to fill in theoretical gaps.

(Tastebud Space) Synesthetic Recipes (Liu, Hockenberry & Selker 2005) is an interface for browsing for food recipes by imagined tastes of food. For example, typing “old, beautiful, desperate, urgent, alive, primal, homey, organic, nutritious, spicy, sweet, moist, aromatic, easy, zen” yields a recipe for “bohemian stew.” With food dishes, tastes, genres, and cultures arranged into a highly connected semantic network, the network approximates a space of taste-for-food. In Synesthetic Recipes, a viewpoint, called a “tastebud,” can be programmed into one of three avatars. As the user browses for food, the avatars constantly emote their likings and dislikings for suggested recipes. An individual’s tastebud can also be acquired through observational learning of what the user types into the search box. This is a minor viewpoint example and will be included in the thesis for completeness.

(Humor Space) Buffalo is a humor robot that suggests jokes it anticipates an individual will find funny. It does so by having
crafted a model of a person’s sense-of-humor relative to the space of jokes. Using a Semantic Sheet representation like What Would They Think?, Buffolo reads an individual’s egocentric texts such as a weblog or email corpus with the goal of extracting (topic, pressure) pairs, much as WWTT extracted (topic, affect) pairs. Pressure is one particular dimension of a full affect measurement. Harkening to psychoanalysis’s hydraulic model of emotions (Freud 1901), an individual’s affective pressure points suggests psychic tensions which need catharsis—humor is a primary means to meet cathartic need (Freud 1905). A major way to structure humor is by culture, since much of one’s embarrassing and tense experiences growing up is shaped by cultural idiosyncrasy, e.g. Asian families and scholastic and work ethic emphasis; overbearing and verbose relatives in Jewish families, narratives of hustling, ghettos, players, and bling in Afro-American culture. Thus, Buffolo senses an individual’s cultural identifications and uses this as a humor viewpoint from which to predict the pleasure of a joke. Buffolo is also a minor viewpoint example and will be included in the thesis for completeness.

2.4 Evaluation
To evaluate the computational theory of point-of-view to be presented in the proposed thesis, I propose both a model-based validation, and task-based evaluation. The model-based validation will evaluate, statistically, how well the machine-learned viewpoint models reflect the reality of an individual’s viewpoint. The task-based evaluation will examine how usable and useful interactive viewpoint artifacts can be in support of various tasks such as recommendation, and supporting constructivist learning of other persons’ perspectives.

Model validation. Among the major viewpoint systems to be discussed in the thesis, space models are computed for perceptual aesthetic space, for opinion space, and for cultural identity/taste space. Location models are computed from individuals’ egocentric texts. I propose to evaluate how well both the space models and location models present accurate pictures of viewpoint spaces and individual viewpoints. I will collect human ratings of the computer-generated models as baselines for computing the accuracy of model acquisition. A pitfall here is self-reporting bias of the subject rating his own subjectivity. I plan to collect appraisal of the subject from friends of the subject. An interesting factor will be the Kappa statistic of inter-rater agreement. The computer models in this thesis can only be expected to perform within the margins of controversy as embodied in the Kappa statistic. I will also evaluate the quality of reactions produced through simulated viewpoints by employing human judges. A couple of these evaluations are already obtained, including a human-rated evaluation of the quality of attitude prediction in What Would They Think?, and a multiple judge evaluation of the signalling efficacy of the aesthetic dimensional space.
Task-based evaluation. The usefulness of viewpoint artifacts speaks to the significance and potential impact of computing viewpoint. User studies will illuminate how well viewpoint artifacts can support a diverse set of tasks such as self-reflection, taste-based recommendation, learning about others, decision support, artistic portrayal, and others. Most of these evaluations have already been performed and obtained, including: a taste-based recommendation task using the viewpoint space of cultural taste (quantitative corpus-based evaluation); a learning about others task using What Would They Think?; an artistic portrayal task in the Aestheticscope. The three evaluations I have already performed were studies conducted with about 20 subjects, and include quantitative analyses. I plan to continue to evaluate from new task angles, so long as my timeline permits it.

3 Contribution
The proposed thesis aspires to be the first comprehensive and computed theory of point-of-view. The theory will be well supported by built viewpoint models for several domains such as aesthetics, cultural identity, and opinions, implementations of automated viewpoint acquisition from readings of text, and implementations of several interactive viewpoint artifacts, which demonstrate broad and significant implications for this line of research. Specifically, I hope to show that

- The slippery notion of point-of-view, well covered in the Semiotics literatures, can be represented, captured, and reified into computational artifacts.
- A point-of-view can be computed most elegantly as an individual’s locations within latent semantic spaces of viewpoint such as OpinionSpace, PerceptualAestheticSpace, and CulturalIdentitySpace.
- A cultural topology can be acquired via culture-mining
- An individual’s point-of-view and the topology of viewpoint spaces can be acquired automatically through psychoanalytic reading.
- Interactive viewpoint artifacts that simulate an individual’s judgments and react to an individual’s actions just-in-time and just-in-context can afford powerful new tools for learning about others, for self-reflection, for inspiration, and for deeper user modeling.

4 Background and Related Work
In the following, I revolve discussion around the major topics treated in this thesis—psychoanalytic reading for viewpoint, viewpoint spaces, point-of-view as ‘locations’ in space, simulating judgment, interactive viewpoint artifacts. For each section I present both computational and non-computational work.

4.1 Psychoanalytic reading for viewpoint
Non-computational work
Psychoanalytic reading means reading to uncover the author of the text. The most important unconscious manifestation of the author in the text is subjective attitude. As discussed in Section 2.1, the semiotic frameworks of Roman Jakobson (1960), JL Austin (1962), and Kaja Silverman (1983) are most central to this work. Jakobson’s theory of six communicative functions implicates six loci of communication and states that all narratives avail primarily one of the loci, be it—sender, message, receiver, context, channel, or code. Psychoanalytic reading is most fruitful over egocentric texts because those texts locate at the sender, and serve the emotive function. Kaja Silverman’s suture technique for psychoanalyzing narratives (1983) suggests that the psychoanalytic reader bears the onus to stitch together lots of disparate impressions it obtains of the author, and suggests affect as a modes of unification. Affect unification was advocated by Friedrich Schleiermacher (1809)—founder of the modern science of hermeneutics, and more recently by George Poulet, who privileges intimacy in reading—“the I who ‘thinks in me’ when I read a book is the I of the one who reads the book” (Poulet 1980, 45).

Psychoanalytic reading can be attempting to either uncover the semi-conscious intent of an utterance, what JL Austin (1962) calls its illocutionary force (e.g. GW Bush’s utterance “You are either with us or you are with the terrorists” has the illocution of a threat), or it can attempt to uncover the collective unconscious underlying each utterance. While reading for the former requires deep story understanding, reading for the latter requires only passive recognition. This thesis computes chiefly the latter because it has access to the space of the collective unconscious as embodied in cultural and linguistic corpora. Reading through the lens of the collective unconscious is termed pejoratively as a hermeneutics of ‘suspicion’ by Paul Ricoeur (1965) but is convalesced by Louise Rosenblatt (1978) as ‘aesthetic reading’, posed in opposition to ‘efferent reading’. ‘Efferent’ means objective reading—reading with the modus operandi to take away something from the text. ‘Aesthetic’ reading means to allow the reader to live through the text—this is the evocative mode of reading which the Aesthetoscope uses to liberate from the text, meanings which come out of the reader, not out of the author. In summary, psychoanalytic reading can either be convergent (romantic, efferent) or divergent (aesthetic, suspicious), and interior (romantic) or exterior (suspicious).

Computational work
Where early AI systems assumed a monolithic model of understanding, coining their work “story understanding” (Winograd 1971; Charniak 1972; Dyer 1983), more recent work reflects a more sophisticated cognitive view, now called computational reading. Moorman and Ram’s (1994) ISAAC reader could focus, attend, and willfully suspend disbelief, AQUA (Ram 1994) could interleave and motivate reading with the asking of questions. Srinivas Narayanan’s KARMA system (1997) reads texts metaphorically, using Petri-nets to understand physical metaphors in text, e.g. “Japan’s economy
stumbled.” ISAAC, AQUA, and KARMA all rely on underlying situation models (Zwaan & Radavansky, 1998), a construct meant to demonstrate rational unification of comprehension. Our work on psychoanalytic reading expands the literature to include affective rather than rational unification.

4.2 Viewpoint spaces

Non-computational work
Viewpoint spaces behave as mathematical tensors do—they attempt to enumerate all possible viewpoints that can be taken within some realm. For some realms, the space of all possible viewpoints tends to be internal to a subject—for example, Jung’s (1921) four modes of psychological function describes the internal space of perceptual aesthetics. For other realms, the space of all possible viewpoints tends to be external—for example, the space of cultural taste can be modeled by mining a representation of the culture, and likewise for opinion space. Roland Barthes and Clifford Geertz presented semiotic representations of culture. Barthes (1964) conceptualized culture as a semiological system of signs and privileges—for example, in Western cultures, ‘rich’ is a privileged sign. In The Interpretation of Cultures (1973), Clifford Geertz posed culture as ‘webs of significance’ and also implicated a person’s internalization of culture—“man is an animal suspended in webs of significance he himself has spun, I take culture to be those webs” (Geertz, 1973: 4-5).

Computational work
The Internet contains many resources such as weblog communities, social networks, recipe corpora, humor corpora, political corpora—these resources are reflections of the cultures of the offline everyday world—thus mining these resources can provide us with working models of viewpoint spaces. The topology of these latent semantic spaces can be inferred through statistical modeling techniques such as Latent Semantic Analysis (Deerwester et al. 1990), Support Vector Machines (Joachims 1998), Multi-Dimensional Scaling (Kruskal & Wish 1978), and the mathematical method of Principle Components Analysis. The culture mining and computational ethnography approach taken in this work parallels the movement of ‘emergent semantics’ (Aberer et al. 2004) which advocates the countervailing view that semantic ontology should be shaped from the ground-up, a posteriori, and in accordance with the natural tendencies of the unstructured data—such a resource is often called a folksonomy when built by humans (e.g. dmoz.org, allrecipes.com).

4.3 Point-of-view as ‘locations’ in space

Non-computational work
Knowing the topology and constitution of viewpoint spaces, I pose an individual’s point-of-view as locations and situations within this space. This view originates in the traditions of psychological situationalism and social constructionism. Par excellence, Pierre
Bourdieu (1984) represents individuals as having a viewpoint system called *habitus*, which has an intersection with the cultural *field*, called the *doxa*. In the epistemology literature, viewpoint is called *perspective* or *subject-position*. Situationalism can be traced to the Sixth Century in the Indian linguistic tradition, can be found in David Hume's (1748) experientialism, and semiotic situationalism was formalized in Jacques Lacan's (1957) theory that ‘the ego is formed out of the other’ (1957) ('other' meaning environment in Lacan's discourse). Applications of situationalism include Georg Simmel's (Levine 1971) finding of self at the intersection of identity fragments (job, church membership, social status); Mihaly Csikszentmihalyi and Eugene Rochberg-Halton's (1981) finding of the self in the 'symbolic environment' of objects in the home; Kevin Murray's (1990) finding of identity formation out of cultural narratives like romantic and comedic stories.

**Computational work**

In AI, it is standard practice to situate the present in priors, as in memory-based reasoning (Stanfill & Waltz 1986), case-based reasoning (Riesbeck & Schank 1989) and reinforcement learning (Kaelbling, Littman & Moore 1996). In user modeling, the user is situated either in stereotypes (Rich 1979) or the user is situated as one vector of actions in a space of many vectors of actions representing the space of other users, i.e. collaborative filtering (Shardanand & Maes 1995). However, reducing a person to but a few categorical attributes lacks the specificity of description necessary to represent one individual, while most behavior modeling is too domain or task-specific and the learned features do not rise the generality of describing an individual as he exists outside applications. Models to be developed in this thesis will locate a person's psychological viewpoint with the specificity of behavioral models, but with the generality of stereotype models.

**4.4 Simulating judgment**

**Non-computational work**

Judgment is chiefly simulated by reference to memory and social priors (in What Would They Think) and by psychological distance (in Taste Fabrics / Interest Map) in the work. The former is supported by reflexive memory theory (Tulving 1983) and imprimer theory (Minsky forthcoming). The latter is supported by H. Montgomery's (1994) “Towards a perspective theory of decision making and judgment” in which Montgomery writes—“Three determinants of perspectives in thinking are identified: (a) the subject, i.e., subject orientation, (b) the object, and (c) psychological distance between subject and object” (Montgomery 1994: abstr.).

Two larger contexts for simulating judgment are mindreading and cognitive perspectivism. On the former, there is recent interest in the cognitive faculty of *mindreading*, *intentionality*, and *theory of mind*. Some theory of mind theorists advocate a simulation approach (Gallese & Goldman 1998), which is sympathetic to this thesis. On the latter, Daniel Dennett (1987) delivers a cognitive explanation of
perspective as the adoption of various stances. For example, a robbery witnessed through the ‘physical stance’ yields physical perceptions like a convenience store opening and a man-object storming out. Witnessed through the ‘design stance’, telic and agentic aspects are illuminated and it is seen that robbers are designed to rob stores, which are designed to carry money. Finally, witnessed through the ‘intentional stance’, it is noticed that the robber is a willful and rational person who robbed this store out of some motivation or habit, and that he is running because he is fleeing from the scene of the crime.

Computational work
The methodology of simulating judgment in this work is reactive in line with the advocacy of BF Skinner’s behaviorism and Rodney Brooks’s (1991) reactive AI, but reactive along affective dimensions, or with respect to connectedness. The chief methods used are thus spreading-activation (Collins & Loftus 1975) and analogical reasoning (Gentner 1983; Fauconnier & Turner 2002).

Other dominant but dissimilar approaches to simulating judgment have been more rational rather than visceral. Dennett’s intentional stance is embodied in the Belief-Desire-Intention model (Georgeff 1998) of agency in the Agents literature. Simulating an agent’s next steps is posed as ‘action selection’ (Maes, 1994) and motivated by goal overloading (Pollack 1992). Other notable rational simulations include Allen Newell and Herbert Simon’s (1963) General Problem Solver, and Newell’s SOAR (1990) cognitive architecture. Other rational-affective hybrid simulations proposed by not built include Marvin Minsky (forthcoming), Push Singh (2005), and Aaron Sloman’s (1981) tiered architectures for minds. Cyc (Lenat 1995), ThoughtTreasure (Mueller 2000) and ConceptNet (Liu & Singh 2004b) represent a large-scale knowledge approach to simulating thought, guided by a topology of thinkables. Cyc and ThoughtTreasure are rational planners, and ConceptNet is reactive and contextual.

4.5 Interactive viewpoint artifacts

Non-computational work
The design of interactive artifacts finds relevance in Human-Computer Interaction studies of synesthetic characters, since by embodying a viewpoint, these artifacts are vulnerable to anthropomorphization. Byron Reeves and Clifford Nass (1996) advocate that anthropomorphic agents obey the same affective contracts as with human-to-human communication. However, this work prefers to de-emphasize the agent metaphor and to emphasize instead the tool metaphor. Here, transparency and trust with the tool must be emphasized (Wheeless & Grotz 1977); thus IdentityMirror, Aesthetiscope, What Would They Think, and Synesthetic Recipes all facilitate introspection and examination of the tool’s capabilities. As a tool, but a very rich tool, an interactive viewpoint artifact facilitates constructivist ‘tinkering’ (Papert & Harel 1991), and ludic or playful activity (Gaver 2001). A user
engaging with the perspective of another person's simulated viewpoint can spur critical self-reflection by presenting 'value fictions' (Dunne 1999).

Computational work
“Software agents” (Maes 1994) are computed embodiments of stereotyped human capabilities, and Pattie Maes explored how they could interactively support human choices such as music selection or browsing the Web, and augment human intelligence (I.A., not A.I.). As stated earlier, I chose not to emphasize the anthropomorphic aspects of the viewpoint artifact so the synesthetic characters literature is not relevant. As a tool, what is relevant are work on interface agents and responsive environments. Bradley Rhodes (Rhodes & Maes 2000) and Henry Lieberman (1997) describe interaction agents that observe user actions such as typing or browsing, and serendipitously and proactively give advice or suggestions. Another line of computational work in Responsive and Reflective Environments (Krueger 1983) investigates how interfaces such as Identity Mirror can engage an individual to ‘perform’ self-reflection.

5 Timeline
I plan to finish refactoring technical implementations and complete all evaluations by the end of January 2006, and the thesis and defence by the end of the spring term 2006.

6 Resources
No additional resources are required, beyond typical access to Media Lab resources, and opportunities to travel to meet with non-local readers.

References


Hegel, G. W. F.: 1807, Phänomenologie des Geistes.


Roget, P.: 1911, Roget's Thesaurus of English Words and Phrases. Retrieved from gutenberg.net/etext/10681


**Candidate Biography**

Hugo Liu is America On-Line Fellow and a dissertation-year Ph.D. candidate in the Media Arts & Sciences program of the MIT School of Architecture and Planning. At the Media Laboratory, Liu is developing a research programme around the computation of point-of-view and aesthetic experience, to culminate in the AAAI-2006 Workshop on Artificial Intelligence for Beauty and Happiness, which he will co-chair. Inspired by literary theory and informed by artificial intelligence computation, Liu’s investigations span the computational modelling of, inter alia, emotion, aesthetics, culture, common sense, gustation, identity, and more recently, he has published on the semantics of happiness and time perception. Liu has published over two dozen articles and spoken extensively around these thematics — his work is recognized with two best paper prizes and has been honoured in New York Times Magazine’s Year in Ideas 2005. He is author of the ConceptNet common sense reasoning research package and the popular MontyLingua natural language understanding package. Liu’s scholarly activities have included service on programme committees of international conferences, reviewerships for major journals and magazines, and keynotes and invited panel presentations. Liu holds bachelor’s and masters’ degrees in Computer Science from MIT.