Anthropomorphism – a Multi-Layered Phenomenon

Per Persson, Jarmo Laaksolahti & Peter Lönnqvist

Swedish Institute of Computer Science (SICS)

Box 1263, 164 29 Kista, Sweden

{perp; jarmo; peterl}@sics.se

Abstract

Anthropomorphism is a way of making sense of complex behavior in the world around us. It is not one process, but involves several – quite independent – levels, e.g., primitive psychology, folk-psychology, social stereotypes and emotional anthropomorphism. This paper explores psychological and anthropological research into the different layers in order to better understand anthropomorphic processes triggered by computer interfaces and systems. Each level is exemplified by existing systems. In conclusion, when referring to anthropomorphism – for instance in relation to an interface or system – we must be clear about which level we indicate, since each has its own characteristics and involves specific types of user expectations.

Introduction

As computer systems become more complex and autonomous – taking on responsibilities users previously had – we will see a shift in users' understanding of computer technology. Whereas the direct manipulation paradigm encouraged users to perceive computers as tools, agent technology will promote users to project anthropomorphic expectations on computer systems. As with any other system displaying complex behavior (e.g., animals, corporations, and machines), users will try to understand and make sense of computer technology through everyday expectations of human life. The interaction 'stance' of users will transform from *action* to *interaction* (Dahlbäck, 1999) from a *mechanistic* stance to an *intentional* stance (Dennett, 1987), from *paradigmatic thought* to *narrative thought* (Bruner, 1986), from *dead tools* to *semi-intelligent* and *socio-communicatory beings*. As designers, we may choose to encourage or discourage such anthropomorphic processes. Irrespective of which, the computer community needs to understand how and under what conditions they work. Anthropomorphism cannot be reduced to observer-independent properties of objects, systems, or creatures. Anthropomorphism is an *experience*, an *understanding* of complex – not necessarily human – patterns of behavior in

the world. Anthropomorphism is a way of simplifying and thereby making sense of the environment by projecting a host of expectations about human life onto aspects of that environment. 'Anthropomorphizing' reality is a *stance*, describing and explaining intricate domains of reality in terms of abstract frameworks of folk-psychology and human life. For instance, users know that computers do not have *beliefs* or *intentions*, but they still attribute such mental states to them in order to make sense of their surface 'behavior' and create 'interaction protocols' with them (cf. Reeves and Nass, 1996). Anthropomorphism has a fundamental 'sense-making' function. Thus, an interface is not anthropomorphic *per se*, but only in so far as it gives rise to anthropomorphic processes in a given user and situation.

Of course, anthropomorphism is not just fantasy and idiosyncratic projections. Some structures or patterns of behavior 'out there' in the 'das-Ding-an-sich' world cue the anthropomorphic stance better than others. Some artifacts 'channel' the anthropomorphic understanding more efficiently than others (see below). Often, however, anthropomorphic understanding requires little support from the environment. In a now classical experiment, Heider & Simmel (1944) showed subjects a film with simple two-dimensional, black triangles and squares moving around on a white surface. When asked to describe what happened in the film, many subjects tended to use anthropomorphic terms and reasoning in order to interpret (and thereby maybe better remember) the events in the film. For instance, the large triangle was described as *chasing* and then *fighting* the smaller one, while the small triangle was *trying* to get away. Although no sophisticated artifact is needed to initiate anthropomorphic thinking, as this example shows, designers need to understand how – and on what levels – anthropomorphism operates in order to produce technology that exploits it in the most efficient way.

Levels of Anthropomorphism

Anthropomorphism resides neither wholly in 'objective reality', nor wholly in the mind of the observer. It arises in the *interaction* between a set of anthropomorphic expectations and external reality (Figure 1). The terms *anthropo-*

morphic or *human-like* are often used as if their meanings are clear and agreed upon.
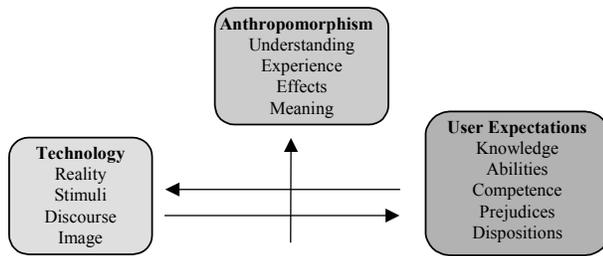

Figure 1. Anthropomorphism emerges in the interaction between technology and user.

We believe that anthropomorphism encompasses a wide range of phenomena, whose characteristics the scientific community lacks a sufficient understanding of. One first step towards such an understanding is to acknowledge that anthropomorphism is not one phenomenon, but in fact exists on different levels. *Anthropomorphism* means different things on different levels.

Primitive Categorization

On the most basic level, people tend to make distinctions between moving and non-moving entities. In particular, self-propelled and autonomous movement – that does not seem to be caused by external forces – attracts attention since it is an indication of 'living matter' (White, 1995). In classical monster films, the convention of having the apparently dead psychopath make a small movement of a part of the body (in close-up), exploits this propensity: something we thought was dead, is now alive and 'intentional' again (cf. also puppets).

Visual appearance of a creature or object seems to be a fundamental trigger of anthropomorphic thinking. Configurations that look like faces and bodies, in reality or in images, have a particular salience to us. Babies seem to be predisposed to look for and fasten their gaze in other people's eyes. Humans appear to be especially equipped to recognize faces and body movements. People anthropomorphize entities that look and move like humans.

Such effects have evolutionary value, which is probably why they have persisted ('Spend particular attention to things that looks like this'). Other primitive categorization can be explained in the same manner. 'Cuteness' – something small and unobtrusive – triggers feelings of 'infantile helplessness' and a tendency to care for and protect (Tan, 1996:160f). Experiences of 'ugliness' and 'beauty' set off an action tendency to avoid or to approach that particular face or body.

The presence of a voice – irrespective of the contents and sounds of that voice – is also a forceful cue to start applying the intentional stance.

Primitive Psychology

Once primitive categorization is in place, people start projecting other kinds of aspects on a given 'living thing'. *Primitive psychology* encompasses expectations about needs, drives, life preservation, sensations and pain. Humans have – probably on the basis of individual experience – 'folk-theories' about how hunger, thirst, sexual drives, and pain work, and the different ways in which they are related. For instance, satisfaction goals such as hunger will disappear if eating, and the ensuing satisfaction will fade over time until hunger will reoccur in a cyclic pattern. Pain will come about if the body is exposed to pressure or beating, and this sensation will be expressed through sound, gestures or facial expressions.

Quite a few systems have tried to exploit (and model) this level of anthropomorphism through animal-like characters – e.g., *Silas* (Blumberg, et al., 1996), *Tristão* (Martinho and Paiva, 1999), *Kismet* (Breazeal, 1999) and *Virtual Petz* (Stern et al., 1998).

Folk-Psychology

Humans also have more sophisticated theories about how the psyche of other people work. On the one hand, people are equipped with a folk-model of the *mind* (Astington et al., 1988; Whiten, 1991; Astington, 1993; Bartsch and Wellman, 1995). Such a model specifies the ways in which *perceptions*, *beliefs*, *goals*, *intentions* and *actions* relate, and how people in everyday talk explain behavior in terms of such 'inner states'. For instance, desires and goals motivate intentions and actions, and beliefs constraint action ('Edi went to the store because she *wanted* an ice cream, and *thought* that the store would be open.'). Such *belief-desire reasoning* slides into place quite early in life, around four years of age, and seems to be fairly universal (Avis and Harris, 1991)

On the other hand, folk-psychology also involves the ability to attribute, and reason about, *emotions* of other people. Within psychology, *appraisal theory* tries to describe how people appraise and ascribe emotion to others (and oneself), e.g., Roseman et al. (1996); Ortony et al. (1988); for overview, see Omdahl (1995). As these investigations show, emotions in a given situation or event are appraised on the basis of a finite set of appraisal parameters, e.g., Is the event (in)consistent with the person's goals?; Is the event caused by the person herself, another person or by circumstance?; Is the event unexpected by the person (contradicting her beliefs)?; Is the event certain or uncertain? Attributing emotions to other people always involves such appraisals: the more appraisal parameters available to us, the more nuanced the attribution.

There have been a couple of efforts to promote the folk-

psychological level of anthropomorphism by supporting this kind of reasoning in consistent ways. The different applications coming out of the OZ project (*Lyotard*, *The Woggles*), model some aspects of primitive psychology as well as folk-psychology (Bates, et al., 1992). *Puppet* specifically deals with supporting children to develop a model of the mind (André, et al. 1999). The behavior of *Sam* not only triggers folk-psychological reasoning in the user, but Sam also has the ability to employ appraisal abilities in understanding a given situation (Elliott, 1992).

Traits

Trait is another level at which to understand an anthropomorphic system. Trait terms are handy ways of summarizing one's impression of a person or a system, and we use such terms extensively in everyday life: *shy*, *optimistic*, *thoughtless*, *stupid*, *extrovert*, *aggressive*, *curious*, *selfish*, *pragmatic*, *nervous*, *idealistic*, or *self-confident*. In contrast to the mental terms in folk-psychology, traits are considered to be more enduring and stable features. They are (attributed) 'dispositions' to behave in a certain way in a certain situation. *Trait* or *personality psychology* is a scientific version of these kinds of everyday attribution (Brody, 1994).

Moreover, traits seem to be short hand terms for complex processes on the folk-psychological level of anthropomorphism. An *optimist* or a *self-confident* person, for example, has a *high control potential* is thus associated with emotions like *frustration*/*disgust* rather than *sadness*/*distress*, or *anger* rather than *dislike*. An *impulsive* or *thoughtless* person downplays *thinking* and *beliefs* in constraining action, and acts primarily on the basis of *desires*. A *dreamer* revels in *desires*, but seldom forms *intentions* or *actions*. A *curious* person is open for *perceptions* to change *beliefs*, whereas a *prejudiced* individual lets his *beliefs* bias *perception*. *Mean* people are without emotions of *regret*, *guilt* and *shame*, which enable them to go on doing bad things to other people.

As these examples show, traits are tightly linked with folk-psychology. *The Inhabited Market Place* (André & Rist, 2000) is one example that specifically models this level of anthropomorphic processes.

Social Roles

Interacting with other people in an everyday socio-cultural environment, often requires expectations about social roles that people will play in a given situation or relation. Understanding a situation in terms of such *social schemas*, is a fundamental dimension among humans, and we should expect people to project such expectations on complex systems as well. A number of social schemas have been proposed by psychologists and sociologists.

One type of social schemas are *occupancy roles* (Taylor & Crocker, 1981:91; Augoustinos & Walker, 1995:39) providing us with normative expectations about the daily activities and standard whereabouts of *doctors*, *waiters*, *police officers*, *scholars*, *chefs*, *farmers*, and *bus drivers*. Such role schemas are often part of *event schemas* or *scripts* (Schank and Abelson, 1977; Bower et al., 1979), and may contain expectations on goals, beliefs, emotions, morals, and behavior of those roles.

Occupancy roles lie close to family role schemas, e.g., *mother*, *father*, *children*, *cousins*, *sisters*, *siblings*, *uncles*, or *lovers*, which contain expectations on how such roles interact with each other on a daily basis.

Third, *social stereotypes* constitute a rich mix of assumptions about other people (Augoustinos and Walker, 1995). They can be conceived of as traits ascribed to a group of people. For instance, in our culture, women are considered to be emotional; bachelors are held to be macho and interested in sexual conquests; doctors and scientists are rational and emotionally neutral; farmers are held to be non-intelligent and slow; and the stereotypical Japanese is considered industrious, polite and clever.

These cognitive expectations about traits and behavior are often mixed together with emotional and moral judgments. Stereotypes contain an *evaluative* 'verdict' of the category of people included in the stereotype, and this is the major difference between stereotypes and other social schemas, e.g., occupancy roles (Tan, 1996:168). Mostly, such evaluations are negative ('Farmers are dim-witted'), but might also be positive or mixed.

In addition, stereotypes contain a *graphical* or *iconographic* dimension. Expectations about traits and personality are often tightly linked to salient external marks, clearly discernable in first-encounters with new people: skin color, hair-color, body size, man/woman, clothing and age (Augoustinos & Walker, 1995: 39f). In first encounter categorizations of another person, this 'visuality' acts as a trigger of stereotype expectations. This is true about real as well as fictional people in visual media. With increased familiarity with a person, visual cues become less important.

Stereotypes are mental representations in individual minds. On a psychological level, people employ stereotypic thinking in evaluating and creating impression in first encounters with other people. Stereotypes, however, also exist on a social level, circulating in mass media and other discursive practices. Discourse – in all of its forms – creates, maintains, transforms and 'naturalizes' stereotypes. Indeed, "most stereotypes are not based upon valid experience but are based on hearsay or images concocted by the mass media or are generated within our heads as ways of justifying our own prejudices" (Aronsson, 1988; cf. also Quinn & Holland, 1987). Visual media, with its clear iconographic quality, seem particularly pertinent in this respect.

Since stereotypes contain an evaluative dimension, stereotyping may cause harm to the person categorized. Moreover, stereotypes influence not only social *cognition*, but also social *action*. In everyday life and in institutionalized and public spheres of social life, stereotype-based action is often profoundly discriminatory and painful for individuals and groups of individuals. Stereotypes are particularly dangerous if they have become 'naturalized and objectified knowledge', seemingly describing 'the way things are'.

Stereotypes, however, also have an instrumental role in social cognition, which is probably the reason why all human cultures have developed systems of them. Just like any other form of schema, social stereotypes are highly functional and cognitively economical since they *simplify* a given set of inputs, and may even predict impending behavior. Stereotypes are frameworks of social expectations, and by projecting these on the environment people do not have to build new models for every new social stimulus. Without social schemas, it would be cognitively impossible and extremely time consuming to understand the social environment. Stereotyping can be an adaptive, shorthand way of dealing with complex events.

As with any other schema, stereotypes are – to use Hutchins's (1980:12) words – "often transparent to those who use them. Once learned, it becomes what one *sees with*, but seldom what one *sees*." Stereotypes can change over time, but they are usually hard to alter. In either case, they often become public issues and subject for intense debate, at least in Western Europe and the United States.

Although some games and computer systems cue stereotypical thinking in the player/user, we know of no examples that systematically and explicitly tries to model this level of anthropomorphism. Furthermore, there are no systems that actively take a critical perspective on stereotypes or encourage the user to become aware of how stereotypes work in social life. We are right now in the process of designing a socio-emotionally rich game environment that addresses precisely these issues (The Kaktus project – http://www.sics.se/humle/projects/kaktus). We intend to use stereotypes as a way for the player to appraise and evaluate characters in the game. We will try to model stereotypes and stereotypic behavior found in mainstream film, theater or commedia del arte (Murray, 1997).

Emotional Anthropomorphism

All levels of anthropomorphism discussed so far involve cognitive expectations that humans project on behaviorally complex systems like human behavior, animals and perhaps even computers (depending on how we design them). However, anthropomorphism also extends into the affective realm.

In the short timeframe, *moral judgments* and *identification* with another person are highly emotional projections.

Stereotypes often contain moral aspects of this sort, but they may also arise out of a situation ('I think she had a right to be angry with him, since he was unfaithful to her'). Categorizations in narratives of good and bad, hero and villain, princess and dragon, includes precisely this kind of affective bonding. In these cases, one does not only understand why another person acts as she does – by attributing folk-psychology, traits and social roles – but one also takes a moral or emotional stance vis-à-vis that individual. In the former case one *aligns* with a character, whereas the latter involves processes of *allegiance* (Smith, 1995). Both interact in complex and dynamic ways, whose details are still unknown by psychology, anthropology, literary and cinema studies.

Long-term emotional bonding includes processes like *friendship* and *love*. Although humans are capable of triggering such responses from other humans, no systems of today (we believe) can accomplish this level of anthropomorphism.

Conclusion

In conclusion, anthropomorphism involves a range of processes, occurring on different levels. When referring to anthropomorphism – for instance in relation to an interface or system – we must be clear about which level we indicate, since each has its own characteristics and involves specific types of user expectations. In fact, the multi-layered schema presented here may act as a tool for analyzing systems with respect to anthropomorphic effects.

References

André, E., Klesen, M., Gebhard, P., Steve, A., et al. (1999) Integrating Models of Personality and Emotions into Life-like Characters, *Workshop on Affect in Interactions Towards a New Generation of Interfaces, Annual Conference of the EC I3 Programme,* Siena, Italy.

André, E. and Rist, T. (2000) Presenting through Performing: On the Use of Multiple Lifelike Characters in Knowledge-based Presentation Systems, *Intelligent User Interfaces,* New Orleans.

Aronsson (1988) *The Social Animal (Fifth Edition)*, New York: W.H. Freeman and Company.

Astington, J., Harris, P. and Olson, D. (Ed.) (1988) *Developing theories of mind*, Cambridge: Cambridge University Press.

Astington, J. W. (1993) *The child's discovery of the mind,* Cambridge, Mass.: Harvard University Press.

Augoustinos, M. and Walker, I. (1995) *Social cognition: an integrated introduction,* London: Sage.

Avis, J. and Harris, P. (1991) Belief-Desire Reasoning

among Baka Children: Evidence for a Universal Conception of the Mind, *Child Development,* (62), pp. 460-67.

Bartsch, K. and Wellman, H. M. (1995) *Children talk about the mind,* New York: Oxford University Press.

Bates, J., Loyall, B. and Reilly, S. (1992) *An Architecture for Action, Emotion, and Social Behavior*, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA.

Blumberg, B., Todd, P. and Maes, P. (1996) No Bad Dogs: Ethological Lessons for Learning in Hamsterdam, *Fourth International Conference on the Simulation of Adaptive Behavior,* Cambridge, MA., MIT Press.

Bower, G., Black, J. and Turner, T. (1979) Scripts in Memory for Texts, *Cognitive Psychology,* (11), pp. 177-220.

Breazeal, C. (1999) Robot in Society: Friend or Appliance?, *Workshop on Emotion-Based Agent Architectures (EBAA), Autonomous Agents '99,* Seattle.

Brody, N. (1994) Traits, in *Encyclopedia of Human Behavior,* : Academic Press, pp. 419-25.

Bruner, J. (1986) *Actual minds, possible worlds,* Cambridge, MA, US: Harvard University Press.

Dahlbäck, N. (1999) Some Suggestions for Expanding the Conceptual Framework for HCI Development and Evaluation, *C. Basic Research Symposium*, Pittsburgh, May 15-20.

Dennett, D. C. (1987) *The intentional stance,* Cambridge, Mass.: MIT Press.

Elliott, C. (1992) *The Affective Reasoner: A process model of emotions in a multi-agent system* Northwestern University Evanston, Illinois.

Heider, F. and Simmel, M. (1944) An Experimental Study of Apparent Behavior, *American Journal of Psychology,* 57(2, April), pp. 243-59.

Hutchins, E. (1980) *Culture and Inference: A Trobriand Case Study,* Cambridge: Harvard University Press.

Martinho, C. and Paiva, A. (1999) Developing Pathematic Agents, *Workshop on Emotion-Based Agent Architectures (EBAA), Autonomous Agents '99,* Seattle.

Murray, J. H. (1997) *Hamlet on the Holodeck : the Future of Narrative in Cyberspace,* New York: Free Press.

Omdahl, B. L. (1995) *Cognitive Appraisal, Emotion, and Empathy,* Mahwah: Lawrence Erlbaum Associates.

Ortony, A., Clore, G. L. and Collins, A. (1988) *The cognitive structure of emotions,* Cambridge: Cambridge University Press.

Quinn, N. and Holland, D. (1987) Culture and Cognition, in *Cultural Models in Language and Thought,* H. Dorothy and Quinn, N. Cambridge: Cambridge University Press, pp. 3-40.

Reeves, B. and Nass, C. (1996) *The Media Equation: How People Treat Computers, Television, and New media Like Real People and Places,* Cambridge: Cambridge University Press.

Roseman, I., Antoniou, A. A. and Jose, P. (1996) Appraisal Determinants of Emotions: Constructing a More Accurate and Comprehensive Theory, *Cognition and Emotion,* 10(3), pp. 241-77.

Schank, R. C. and Abelson, R. P. (1977) *Scripts, Plans, Goals, and Understanding : An Inquiry into Human Knowledge Structures,* Hillsdale, N.J.: L. Erlbaum Associates.

Smith, M. (1995) *Engaging characters: fiction, emotion, and the cinema,* Oxford: Clarendon Press.

Stern, A., Frank, A. and Resner, B. (1998) Virtual Petz: A Hybrid Approach to Creating Autonomous, Lifelike Dogz and Catz, *The Second International Conference on Autonomous Agents,* Menlo Park, AAAI Press.

Tan, E. S. (1996) *Emotion and the structure of narrative film : film as an emotion machine,* Mahwah, N.J.: Erlbaum.

Taylor, S. and Crocker, J. (1981) Schematic Bases of Social Information Processes, in *Social Cognition,* T. Higgins, Herman, P. and Zanna, M. Hillsdale, N.J.: L. Erlbaum Associates, pp. 89-134.

White, P. (1995) *The Understanding of Causation and the Production of Action. From Infancy to Childhood,* Hove: Lawrence Erlbaum.

Whiten, A. (Ed.) (1991) *Natural theories of mind: Evolution, development and simulation of everyday mindreading*, Oxford, England: Basil Blackwell, Inc.