Using Embodied Allegories to Design Gesture Suites for Human-Data Interaction

Francesco Cafaro
Department of Computer Science
University of Illinois at Chicago
fcafar2@uic.edu

ABSTRACT
Human-Data Interaction (HDI) systems can be defined as technologies that use embodied interaction to facilitate the users' exploration of rich datasets. As the design of gestures for Whole-Body Interaction is often based on an uninformed trial and error approach, I propose the use of Johnson's *embodied schemata*, extended with *allegories*, to inform the design of suites of gestures for HDI systems. This approach involves: (1) the identification of embodied allegories to encourage data exploration, (2) a study of the allegorical relation between input and output, and (3) an analysis of the implication of the (social) space.

Author Keywords Gestures, Metaphors, Human-Data Interaction, Embodied User Interface, Informal Learning.

ACM Classification Keywords H.5.2 [Information Interfaces and Presentation]: User Interfaces: Input Devices and Strategies, Interaction Styles.

General Terms Design, Human Factors.

INTRODUCTION
In the last year I have been involved in the development of CoCensus [10], an interactive system designed to engage museum visitors in inquiries with geo-referenced census data (Figure 1).

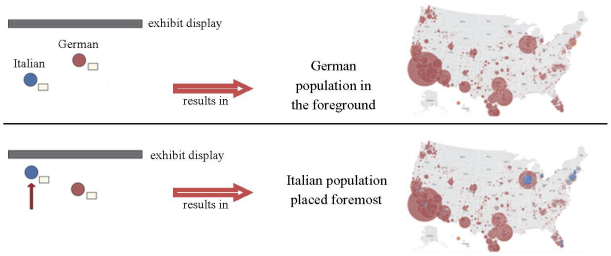

Figure 1. Collaborative data exploration with CoCensus. When visitors approach the display, their data becomes more prominent (via opacity and z-ordering); their body movements (from jumping to arm gestures) cause the subset to "mirror" their actions (momentarily pushing the "bubbles" away).

When entering the museum, a visitor picks an ancestry category (e.g., German), and receives an RFID tag; after approaching an interactive area, she/he sees "her/his" data subset (e.g., the number of German immigrants, represented as scaled "bubbles") on a shared display, along with the data subsets of the other visitors in the room. Each visitor is both *tracked* (using a Microsoft Kinect™ camera) and *identified* (with a passive RFID system).

CoCensus is only one example of the novel interaction possibilities that are now available to the masses thanks to some recent off-the-shelf sensors, such as Nintendo Wii and Microsoft Kinect. The potential of these commercial technologies goes far beyond the gaming consoles they are designed for. In 1986, Buxton imagined a future archeologist digging up a current personal computer and thinking that the user had one dominant hand, one eye, limited hearing and no legs [5]. The wide availability of body sensing technologies may drastically change this scenario, and open up richer ways of interacting, which are closer to the idea of "*Whole Body Interaction*" [6].

Identifying a set of gestures to be used in whole-body interaction systems is, however, still a challenge. While a number of technologies (such as [13] and [14]) has been developed to design personalized gestures with multi-touch devices, the users of a whole-body interactive system need to be instructed on the gestures and movements that the system can detect [7]. Furthermore, the mapping from the movement of body parts to a set of gestures is often based on an uninformed trial and error approach [7].

In 1980, Lakoff and Johnson introduced the idea that metaphors are not only a device of poetic imagination, they play a fundamental role in the way we experience and understand our world and ourselves [12]. Lakoff and Johnson's intuition is further elaborated in [11], with the definition of "*embodied schemata*" and their role in the theory of meaning: they are "structures for organizing our experience and comprehension." For instance, we experience "*containment*" and boundedness everyday (e.g., when we pour water in a cup, or when we move in/out of rooms). An "*embodied schema*" is such a recurrent pattern, shape and regularity in our daily *body experience*; it emerges as a meaningful structure for our movement through space and manipulation of objects.

Embodied schemata may be metaphorically projected, creating common structures to guide our understanding and reasoning. For example, "*balance*" is a bodily experience that a baby learns when she/he stands, falls to the floor, and

UbiComp'12, September 5-8, 2012, Pittsburgh, USA.
ACM 978-1-4503-1224-0/12/09.

stands again, until she/he understands how to keep a balanced erect posture. This schema rapidly evolves metaphorically, so that we can unconsciously apply the concept of "balance" to other domains, such as visual arts and architecture [11].

In the following sections, I will introduce related works that explore the use of embodied schemata to the design of embodied and tangible systems. I will then show how schemata might be adapted for needs specific to Human-Data Interaction, and I will present an outline of my planned research in this area.

RELATED WORKS

In Springboard [3], the concept of "balance" is explored in the domain of social justice, using a whole-body interactive environment; Antle's work is one of the first attempts to use Johnson's theoretical framework to inform the design of whole-body interactive systems. Sound Maker ([1] and [2]) also explores the use of "balance" and the metaphor "music is a body movement", in the design of a system that connects body movements to sounds, in order to teach abstract music concepts such as "tempo."

Embodied schemata (and the metaphors that unconsciously develop from them) have also been used in the design of tangibles. For instance, [9] explores the use of embodied schemata to design a "Tangible Memory Box", where each item in the box is linked to a collection of digital media; Moving Sounds Tangibles [4] introduces a framework for grounding the design of tangibles on embodied schemata.

HUMAN-DATA INTERACTION (HDI)

Making sense of large datasets may be problematic. Let's imagine printing a map of the United States with all the data from the last Census (race, age, population density, housing characteristics, etc.). With so much information, this map would be useless, as some data would occlude others. The user's experience would be totally different, instead, with an interactive map where each person can select and control the variables she/he wants to explore.

In this paper, I use the term *Human-Data Interaction* (HDI) to denote the problem of delivering personalized, context-aware, and understandable data from big datasets. I also refer to *HDI systems* as those technologies that approach HDI within the *Whole Body Interaction* framework [6]. CoCensus [10] is an example of such systems.

I believe HDI should be *embodied*: people want to *explore*, not to *analyze*, data, especially in informal learning settings. "Analysis" implies already having some idea or goal in mind; data exploration is very different – museum visitors are not trying to reach a specific goal, but rather, to come to understand a data set. Also, part of the learning that will occur is in the emergent discussions that visitors have with one another about "their" data sets, when they collaboratively *explore* "their" own data in a common interaction space.

CHALLENGES IN DESIGNING GESTURES FOR HDI

How can we design gestures to interact with HDI systems? We cannot refer to the current attempts of grounding the design of interactive systems on Johnson's embodied schemata. Antle's 1:1 conjunctions of action:effect are more embodied "similes" than embodied metaphors. Similes are a more precise form of metaphor that highlight a particular aspect of the referent domain as being salient for the purpose of comparison, such as "cheeks red like a rose". While Antle makes a very direct relation between the characteristics of the gesture (e.g. running fast) and an abstract concept that she wants to *teach* (e.g. fast tempo), HDI gestures are meant to facilitate the *exploration* of a rich dataset. In order to do so, we are designing embodied "*allegories*", which are extended metaphors where there are multiple symbolic actions and representations used as a unified suite.

Embodied Allegories to encourage data exploration

In CoCensus, the purpose of jumping is not to use some bubbles to teach the abstract concepts of free fall to museum visitors. Walking back and forth (which affects the transparency of the bubbles) is not meant to teach the idea of alpha-value. These gestures are designed to facilitate the users' collaborative exploration of large datasets. CoCensus is based on the *allegory* of an "imperfect mirror": visitors approaching the exhibit should see "their" data on-screen reflect their body movement. This is an *allegory*, not just a metaphor, since it supports a number of actions (forward-back, left-right, up-down), which are all individualization of the "*reflection*" that one would expect to come from a "mirror". An embodied metaphor would just be one of these. Can Johnson's framework be extended with *allegories* to ground the design of HDI gestures?

Allegorical relation between input and output

As observed in [2], metaphors have been usually adopted in conjunction with embodied interaction both for providing input and to facilitate the understanding of the system output. What happens when this mapping between input and output is not as linear as in the case of "fast" movements mapped to "fast" tempo in SoundMaker [2]? When the user controls the transparency of the bubbles by walking in the interaction space, input (the distance from the screen) and output (the variations in the alpha-value) do not seem to be so strictly related. Can we identify some *allegories* that can help users' understanding of the system's feedback?

Implications of the (social) space

Informal learning settings, such as museums, are different from laboratories and classrooms. People may have different cultural backgrounds; there are also differences in age and interests. Does the social space have implications on the gestures that people are willing to perform? Is this subset of gestures based on different *allegories*?

Furthermore, visitors may approach interactive systems in different ways: they may look at them from a distance,

watch other people play with them, or get closer to them in order to have a more personalized interaction. These interaction patterns can be described using the *theory of proxemics* and, specifically, the notion of four "*zones of personal space*" introduced by Hall in [8]. As being at a different distance from the system may involve having different interaction possibilities, I believe that people may use different subsets of gestures. Are there some gestures that are never/always used at a given distance from the screen?

METHODOLOGY

Over the course of three semesters, I will conduct a series of in-situ studies in historic and science museums, using different interfaces to explore the data. My work will be organized in three phases.

Phase 1: Identification of a library of gestures for HDI

Can common embodied schemata and allegories be identified?

I will show a number of animations on the data visualized on the screen to museum visitors. For instance, when using the CoCensus interface, some animations will involve the scaled bubbles (momentarily pushing them left/right/up/down), the evolution of the data along a timeline, etc. I will then ask each individual visitor which gesture she/he would like to do to make the system perform each animation. I will ask visitors to think aloud and explain why they recommend a specific gesture.

Data: I will analyze the use of the language, looking for communalities, embodied schemata and allegories or other figures (metaphors, metonyms, personifications, etc.)

Expected outcome: I believe that an analysis of visitors' recommendations will reveal common schemata (such as in/out, path, counterforce), and that some of the gestures will be inspired on allegorical reasoning.

Is it beneficial to use "allegories" to design gestures?

I will design a series of gestures based on the common embodied schemata and *allegories* identified in the previous step. I will also implement different gestures, based on disconnected metaphors, to trigger the same set of animations. I will ask museum visitors to interact with both versions of the system and to rate their experience. I will investigate how users will progress through the four levels of interaction introduced in [15]: (1) realizing that the system is interactive; (2) identifying a personal role in producing an effect to the shared screen (i.e., recognizing which data subsets they personally affect); (3) exploring bounds of producing the effect; (4) instrumental operation with the system/control of the state of the shared space.

Data: I will evaluate the time required by the users to progress through the levels of interaction, under both conditions. I will compare the ratings assigned by visitors to each condition. I will collect comments and recommendations to verify if users are still reasoning using the previously-identified embodied schemata.

Expected outcome: I suspect that users will move faster through the four levels of interaction when the gestures are grounded on common *allegories*: they will discover related actions faster, and move through them faster, than with the use of disconnected metaphors. I also suspect that they will enjoy this kind of gestures more.

Phase 2: Identification of metaphors across input and output in HDI systems

Are common embodied schemata/allegories used by people when describing the data visualization?

I will provide different static representation of the same data to people and ask them to describe what they see. When using the CoCensus interface, for example, two maps (one with and one without county borders) will be displayed, to investigate the use of the in/out schemata.

Data: I will analyze the use of the language, looking for communalities, embodied schemata and allegories.

Expected outcome: I believe that visitors will refer to common embodied schemata and will use an allegorical language when describing the data.

Are there recurrent allegories used to describe both the gesture and the visualization?

I will compare the results of the visualization study described above with the results of the user study performed during the first phase of my research. I will try to identify a catalog of <gesture, visualization> pairs, which are grounded on the same *allegories*.

Data: I will analyze the language used by visitors to describe their interaction with the system (i.e. the gestures they perform) and the system output (i.e. the data visualization), looking for common allegories across gestures and visualization.

Expected outcome: I believe that people will tend to use similar patterns, both to figure out how to provide input to the system and to understand the output of the system.

Is it beneficial to use a set of <gesture, visualization> which is based on common allegories?

I will implement a version of the system based on the previously identified pairings of <gesture, visualization>, and one where gestures and visualization are independent. I will ask museum visitors to interact with both versions of the system and to rate their experience. I will investigate how users will progress through the four levels of interaction introduced in [15].

Data: I will evaluate the time required by users to move through the four levels of interaction, under both condition. I will compare the ratings assigned by visitors to each condition. I will collect comments and recommendations to

verify that users are still reasoning using the common allegories that I have identified.

Expected outcome: I suspect that users will move faster through the four levels of interaction when there is "allegorical coherence" between gesture and visualization. It is unknown whether they will enjoy the system more.

Phase 3: Implications of the (social) space

Does the social space have implications on the gestures that people may be willing to perform?

At the end of their interaction with the system, I will ask visitors if they would have used different gestures to perform some of the tasks when in the museum or at home.

Data: I will analyze the use of the language, looking for differences in the embodied schemata and allegories when people refer to a social context (i.e. a museum) or to a private one (e.g. their homes).

Expected outcome: unsure.

Is there a relation between gestures and zones of interaction?

Across the several testing iterations, I will observe how people move within the interaction space and which gestures they perform when they are within a given zone of interaction.

Data: I will record a log file with the gesture recognized by the system, the distance of the users from the screen, and the time since when the interaction started.

Expected outcome: I suspect that, when a visitor will approach the screen, she/he will start trying more gestures.

CURRENT STATUS

I have implemented the software infrastructure that intelligently merges the inputs from a camera tracking system and an RFID system. I have also being involved in the design and implementation of a custom visualization for geographically-referenced data. I have run multiple user studies of the CoCensus prototype at the Jane Adam's Hull House, an historic museum and cultural center in Chicago. I am currently starting the first phase of my study, which should lead to the identification of a library of gestures, grounded on embodied schemata, and which will set the stage for the research plan described in the previous section.

PRELIMINARY RESULTS

I have been a key collaborator on a formative user study, with nine museum visitors, designed to evaluate different animation styles for CoCensus. Each time the user was jumping, a different animation was triggered: bubbles (representing the density of population) on the map moved up and down, either following the same patterns or a specific physics model (for instance the biggest ones, i.e. those with the "biggest mass", were slower in rising up). Visitors expressed a general preference for the animations based on a physics model; in other terms, users preferred a system where visualizations and gestures were rooted in the same embodied allegories (the physics model), rather than on distinct embodied metaphors.

CONTRIBUTION

The contribution of this work will be an approach, based on allegories, to design gesture suites for embodied interaction systems. This is likely to produce gestures that can be learned and discovered more easily than in the case of disconnected metaphors. This work will also produce a catalog of gestures for HDI systems.

REFERENCES

1. Antle, A.N., et al. Designing to support reasoned imagination through embodied metaphor. *Proceeding of CC 09*, (2009), 275.
2. Antle, A.N., Droumeva, M., and Corness, G. Playing with The Sound Maker : Do Embodied Metaphors Help Children Learn ? *Proceedings of IDC 08*, (2008).
3. Antle, A.N. Springboard : Exploring Embodiment , Balance and Social Justice. *Social Justice A Journal Of Crime Conflict And World Order*, (2009), 3961-3966.
4. Bakker, S., Antle, A.N., and Den Hoven, E. Van. Embodied metaphors in tangible interaction design. *Personal Ubiquitous Comput. 16*, 4 (2012), 433-449.
5. Buxton, W. and Myers, B. A Study in Two-Handed Input. *CHI 86*, ACM Press (1986), 321-326.
6. England, D. Whole Body Interaction: An Introduction. In *Whole Body Interaction* Springer London, 2011.
7. Fothergill, S., Mentis, H., Kohli, P., and Nowozin, S. Instructing people for training gestural interactive systems. *Proceedings of CHI 2012*, ACM (2012).
8. Hall, E.T. *The Hidden Dimension*. Doubleday, Garden City, N.Y., 1966.
9. Hurtienne, J. and Israel, J.H. Image schemas and their metaphorical extensions. *Tangible and Embedded Interaction*, ACM Press (2007), 127.
10. Roberts, J.A., Lyons, L., Radinsky, J., Cafaro, F. *Connecting Visitors to Exhibits through Design:* Exploring United States census data with CoCensus. Will appear in the *Proceedings of ICLS 2012*.
11. Johnson, M. *The Body in the Mind*. The University of Chicago Press, 1987.
12. Lakoff, G. and Johnson, M. *Metaphors we live by*. University of Chicago Press, 2003.
13. Lü, H. and Li, Y. Gesture coder: a tool for programming multi-touch gestures by demonstration. *Proceedings of CHI 2012*, ACM (2012), 2875-2884.
14. Ouyang, T. and Li, Y. Bootstrapping personal gesture shortcuts with the wisdom of the crowd and handwriting recognition. *Proceedings of CHI 2012*, 2895-2904.
15. Williams, A., et al. From interaction to participation: Configuring space through embodied interaction. *UbiComp 2005.*