Objects, Containers, Gestures, and Manipulations: Universal Foundational Metaphors of Natural User Interfaces

Ron George
Ron George Design
Kirkland, WA
ron@rongeorge.com

Joshua Blake
InfoStrat
4301 Connecticut Ave, Suite 451
Washington, DC 20008 USA
joshb@infostrat.com

CHI 2010, April 10–15, 2010, Atlanta, Georgia, USA.
ACM 978-1-60558-930-5/10/04.

Abstract
In this paper, we propose Objects, Containers, Gestures, and Manipulations (OCGM, pronounced like Occam's Razor) as universal foundational metaphors of Natural User Interfaces. We compare OCGM to existing paradigms using SRK behavior classification and early childhood cognitive development, and justify the "universal" and "foundational" descriptors based upon cognitive linguistics and universal grammar. If adopted, OCGM would significantly improve the conceptual understanding of NUIs by developers and designers and ultimately result in better NUI applications.

Keywords
Natural User Interfaces, Objects, Containers, Gestures, Manipulations, interaction design, metaphor

ACM Classification Keywords
H.1.2. [Models and Principles]: User/Machine Systems---Software psychology; H.5.2 [INFORMATION INTERFACES AND PRESENTATION (e.g., HCI)]: User Interfaces---Theory and methods

General Terms
Design, Human factors, Theory

Introduction

The phrase "Natural User Interface" (NUI) has entered the mainstream consciousness. Non-technical publications such as Wall Street Journal [1] now discuss NUIs, and hundreds of software vendors and design firms have started to include Natural User Interface offerings for devices such as Microsoft Surface [2]. Research on multi-touch and other input modalities has been on-going since the 1980's [3], but there is little published theory or specific guidance on the design and development of Natural User Interfaces. The majority of designers and developers working on NUIs do not have extensive research backgrounds in this field and therefore struggle with unclear ideas and mixed concepts. Consequently, many current NUI software projects have significant usability issues.

One of the points of confusion is the meaning and implication of the word "natural" in the phrase "Natural User Interface." In this context, natural is often used interchangeably with intuitive; however, this word is equally ambiguous and does not help with comprehension of the concepts.

We adopt the definition enunciated by Buxton [4]: An interface is natural if it "exploits skills that we have acquired through a lifetime of living in this word." Buxton elaborates and categorizes skills into innate and learned. The implication is that NUIs will exploit a different skill-set than existing interfaces.

Skills, Rules, & Knowledge Framework

We evaluated the characteristics of existing and proposed interfaces in terms of the skills, rules, and knowledge (SRK) framework [5]. The predominant interface paradigm for the last two decades is the Graphical User Interface (GUI), and the primary interaction style for GUIs can be described as Windows, Icons, Menus, and Pointing devices, or WIMP [6].

In the context of SRK, users interacting with WIMP-style interfaces use rules-based or knowledge-based behaviors, depending upon the technical background of the user. WIMP interfaces have complex metaphors and behaviors that require a relative high cognitive load to use. In some circumstances, such as while driving or overseeing industrial equipment, this cognitive load can lead to unsafe conditions, and has motivated alternative values for GUI interface design [7]. Setting aside unsafe situations, high cognitive load is generally undesirable to common users, and can prevent people with cognitive difficulties or limited mental faculties from being productive [8].

We contend that the limitations of GUIs and in particular the WIMP metaphor restrict our ability to reduce cognitive load beyond a certain threshold. In order to go further, we must identify the various innate abilities of the human brain and design interfaces which take advantage of those natural skills in the appropriate contexts. The ultimate goal is to discover the human-computer interaction principles that produce interface designs which maximize skill-based behaviors.

OCGM are Metaphors

To this end, we propose a new metaphor for Natural User Interfaces: Objects, Containers, Gestures, and Manipulations (OCGM). The acronym can be pronounced Occam, as in Occam's Razor. While performing a broad survey of interaction designs across different device types, co-author George identified a few common patterns and developed the concepts

behind OCGM. [10]. Co-author Blake suggested the acronym order and linked the concept to Occam's Razor [11]. As we will see, the connection to Occam's Razor is apt, as OCGM is based upon the simplest human interactions and our earliest thought patterns.

GUIs commonly use the desktop metaphor to help users understand how to interact with their computers. Metaphors, at their core, help us understand a complex idea from the perspective of a more familiar idea [20], and the same is true of OCGM. Objects are metaphors for units of content or data. Containers are metaphors for the relationships between content. Gestures are metaphors for discrete, indirect, intelligent interaction, and manipulations are metaphors for continuous, direct, environmental interaction.

To better distinguish between Objects, Containers, Gestures, and Manipulations, we can categorize each metaphor into one of four groups, shown in figure 1.

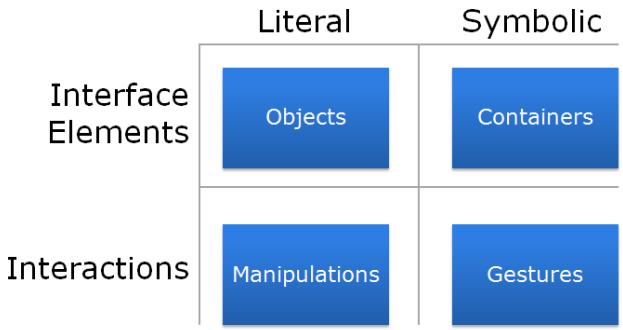

Figure 1. Simple categories reveal the conceptual symmetry of OCGM.

Development of Interaction Skills

In order to determine which skills are used with WIMP and OCGM, we must study the human interaction aspect of HCI. Taking a cue from Piaget [9], we examined the origins of human interactions and early childhood cognitive development as applied to HCI. We evaluated the WIMP metaphor and compared it to the proposed OCGM metaphor.

The cognitive skills required to understand and operate WIMP interfaces are initially developed during or after pre-school age. Hourcade's study comparing four and five year olds to adults concluded the children's poor performance on pointing tasks with mice justifies specialized interface designs [12]. Further, Tolar's study on two to five year olds concluded that the mental capability to judge the meaning of icons did not develop until the age of three and a half [13]. Reading skills necessary for operating typical GUI menus also are not developed until school age.

In contrast to WIMP, the cognitive skills required for OCGM interfaces are developed significantly earlier, nominally by nine months old. Perceptual object categorization has been demonstrated in infants as young as nine months old, with robust object grouping skills by eighteen months [14]. The ability to recognize containment relationships has been demonstrated in six-month olds [15]. Infants develop the ability to recognize and perform gestures to communicate desires between nine and twelve months [16] and can manipulate objects younger than six months [17]. Because OCGM's requisite cognitive skills are developed very early, they are innate and natural. Thus, interfaces using OCGM will have minimal cognitive loading and use skills-based behaviors.

OCGM are Foundational Metaphors

OCGM can also be related to the field of cognitive linguistics, which describes several basic image-schemas[1] which are prerequisites for higher-level image-schemas. For example, establishment of the basic CONTAINER image-schema is required before use of the dependent EMPTY-FULL and EXCESS image-schemas [22]. Santibáñez argues that the three basic image-schemas, upon which all other image-schemas depend, are OBJECT, CONTAINER, and PATH [23]. These fundamental image-schemas coincide with OCGM, with PATH relating to both gestures and manipulations.

Considering the link to the basic image-schemas, OCGM can be thought of as the *foundational* metaphors of all human-computer interaction. OCGM lays a foundation of abstract metaphors that supports many layers of concrete metaphors. Application of cognitive linguistics theories to HCI is justified because the meaning of both speech and gesture is processed by the same neural system [24].

OCGM are Universal Foundational Metaphors

Chomsky studied cognitive development to develop fundamental theories of linguistics such as universal grammar [19]. OCGM and universal grammar have similar roles in their respective fields in that both seek to explain common patterns across many domains. In the case of HCI, OCGM are *universal* foundational metaphors and all human-computer interactions can be described in terms of OCGM, including diverse form factors (desktop, mobile, TV) and modalities (multi-touch, voice, motion sensing, and even mouse.)

[1] Image-schemas are recurring structures in human cognition which represent basic patterns of concepts and understanding [18]. Image-schemas are formed based upon our interaction with the environment.

Summary and Future Work

We proposed Objects, Containers, Gestures, and Manipulations as universal foundational metaphors in the NUI paradigm. OCGM provides a pattern that will lead designers and developers to create more natural interfaces, while being flexible enough to support the implementation of context- and content-appropriate concrete metaphors. We recommend Reality-Based Interactions (RBI), proposed by Jacob et al. [21], as a starting point for designing these concrete metaphors. RBI's themes of Naïve Physics, Body Awareness & Skills, Environmental Awareness & Skills, and Social Awareness & Skills imply OCGM-compatible concrete metaphors that use skill-based behaviors.

While we are confident in the theoretical and conceptual justification for OCGM, there is still work to be done before it can be effectively used in the industry. Specific design and development techniques based upon OCGM should be explored. The effectiveness of both OCGM-based techniques and the resulting interfaces should be quantitatively measured. Finally, succinct and effective documentation of OCGM and related concepts is necessary to gain mind-share in the industry.

Citations

[1] Wingfield, N. "Body in Motion: CES to Showcase Touch Gizmos" Wall Street Journal, January 5, 2010

[2] Kirk, J. "Microsoft Expands Surface to Europe, Middle East" PC World, March 2, 2009

[3] Buxton, B. Multi-Touch Systems that I Have Known and Loved.

http://www.billbuxton.com/multitouchOverview.html. 2007.

[4] Buxton, B. CES 2010: NUI with Bill Buxton. http://channel9.msdn.com/posts/LarryLarsen/CES-2010-NUI-with-Bill-Buxton/. 2010.

[5] Rasmussen, J. Skills, rules, knowledge; signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man and Cybernetics*, 13, 257-266. 1983.

[6] Andries van Dam: *Post-WIMP User Interfaces*. In: *Communications of the ACM*, 40(2) (February 1997), pp. 63-67.

[7] Vicente, K. J. & Rasmussen, J. Ecological Interface Design: Theoretical foundations. *IEEE Transactions on Systems, Man and Cybernetics*, 22, 589-606. 1992.

[8] Gregor, P., & Dickinson, A. Cognitive difficulties and access to information systems - an interaction design perspective. *SigAccess Newsletter.* http://www.sigaccess.org/community/newsletter/september_2005/sep05_all.pdf#page=61. September 2005.

[9] Piaget, J. *Genetic epistemology.* Columbia University Press. 1970.

[10] George, R. OCGM (pronounced Occam['s Razor]) is the replacement for WIMP. http://blog.rongeorge.com/design/ocgm-pronounced-occams-razor-is-the-replacement-for-wimp/. December 28, 2009.

[11] Blake, J. WIMP is to GUI as OCGM (Occam) is to NUI. http://nui.joshland.org/2009/12/wimp-is-to-gui-as-ocgm-occam-is-to-nui.html. December 28, 2009

[12] Hourcade, J., Bederson, B., Druin, A., and Guimbretière, F. Differences in pointing task performance between preschool children and adults using mice. ACM Transactions on Computer-Human Interaction 11(4), December 2004, pp 357-386.

[13] Tolar T, Lederberg A, Gokhale S, Tomasello M. The development of the ability to recognize the meaning of iconic signs. Journal of Deaf Studies and Deaf Education. 2008. 13:71–86.

[14] Gopnik, A., Meltzoff, A. N. The development of categorization in the second year and its relation to other cognitive and linguistic developments. *Child Development*. 1987. 58, 1523-1531.

[15] Casasola M, Cohen LB, Chiarello E. Six-month-old infants' categorization of containment spatial relations. Child Dev. 2003;74:679–693.

[16] Woodward, A.L., & Guajardo, J.J. (2002). Infants' understanding of the point gesture as an object-directed action. Cognitive Development, 17, 1061–1084.

[17] Johnson, C.P., & Blasco, P.A. (1997). Infant growth and development. Pediatrics in Review, 18, 224–242.

[18] Johnson, Mark. *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason*, University of Chicago. 1987.

[19] Chomsky, Noam Lectures on Government and Binding. Dordrecht, Holland: Foris Publications. 1981.

[20] Szabó, K. Metaphors and the User Interface. http://www.katalinszabo.com/metaphor.htm. 1995.

[21] Jacob, R. J. K., Girouard, A., Hirshfield, L. M., Horn, M. S., Shaer, O., Treacy, E. S., and Zigelbaum, J. Reality-Based Interaction: A Framework for Post-WIMP Interfaces. *Proc. CHI 2008*, ACM Press. 2008.

[22] Peña, S. A Cognitive Approach to the Image-Schematic Component in the Metaphorical Expression of Emotions in English. Unpublished Ph.D. Dissertation, University of La Rioja.2000.

[23] Santibáñez, F. The Object Image Schema and Other Dependent Schemas. Atlantis, 24, 183-201. 2002.

[24] Jiang Xu, Patrick J. Gannon, Karen Emmorey, Jason F. Smith, and Allen R. Braun. Symbolic gestures and spoken language are processed by a common neural system. PNAS 2009 106: 20664-20669.