SIMBORGS: TOWARDS THE BUILDING OF SIMULATED HUMAN USERS FOR INTERACTIVE SYSTEM DESIGN

Wayne D. Gray & Michael J. Schoelles
Rensselaer Polytechnic Institute
[grayw/schoem]@rpi.edu

We propose an approach to the cognitive engineering of *integrated task environments* by the use of *simulated cyborgs* (simBorgs). SimBorgs combine high-fidelity computational cognitive models with low-fidelity artificial intelligence (AI) based reasoning components. This combination of cognitive modeling with AI enables the creation of intelligent agents, simBorgs, that will work tirelessly to perform usability testing on various combinations of tasks and interfaces.

For a simulated human user to be useful in testing interface designs how much of human behavior must be simulated in a cognitively plausible manner? For those interested in interface design, the problem is that task knowledge and even cognitive processes (Ericsson & Kintsch, 1995) are very task-specific. Do we need a cognitively plausible model of an expert architect to evaluate the interface of an architectural CAD-CAM system? Do we need a cognitive model of an Air Traffic Controller to evaluate interface features of ATC software?

It is certainly the case that if one is not interested in cognitive fidelity, it is often simpler to model human performance by a mathematical equation, statistical techniques such as multiple regression, or an AI-system than it is to build a high-fidelity computational cognitive model. For example, Deep Blue beat Gary Kasparov at Chess. However, Deep Blue was not a model of human chess playing expertise. Designing an AI system that will beat the world's best human at chess has been done. Designing a computational cognitive model that will play world-class chess in the same manner as the best human plays chess has not been done.

We argue that it may be possible to finesse the expertise issue by turning the reasoning component of a model over to a *black-box module* that makes few, if any, claims to the cognitive fidelity of its processing. Whether this module uses machine-reasoning algorithms, statistic models, or other formalisms will depend on the needs of the particular project. In effect, we are advocating the building of simulated Cyborgs (simBorgs).

SimBorgs and a Three-Tier Architecture for Interactive System Design

Cyborgs are science-fiction creations that are part human and part machine. We see the creation of simulated cyborgs as the solution to building models that interact with a variety of software tools in a manner that requires domain expertise. Like the sci-fi cyborgs, our simBorgs combine human and machine components. Our simBorgs consist of high-fidelity models of human interactive behavior, knowledge schemas derived from expert humans as they perform realistic tasks, and a black-box module.

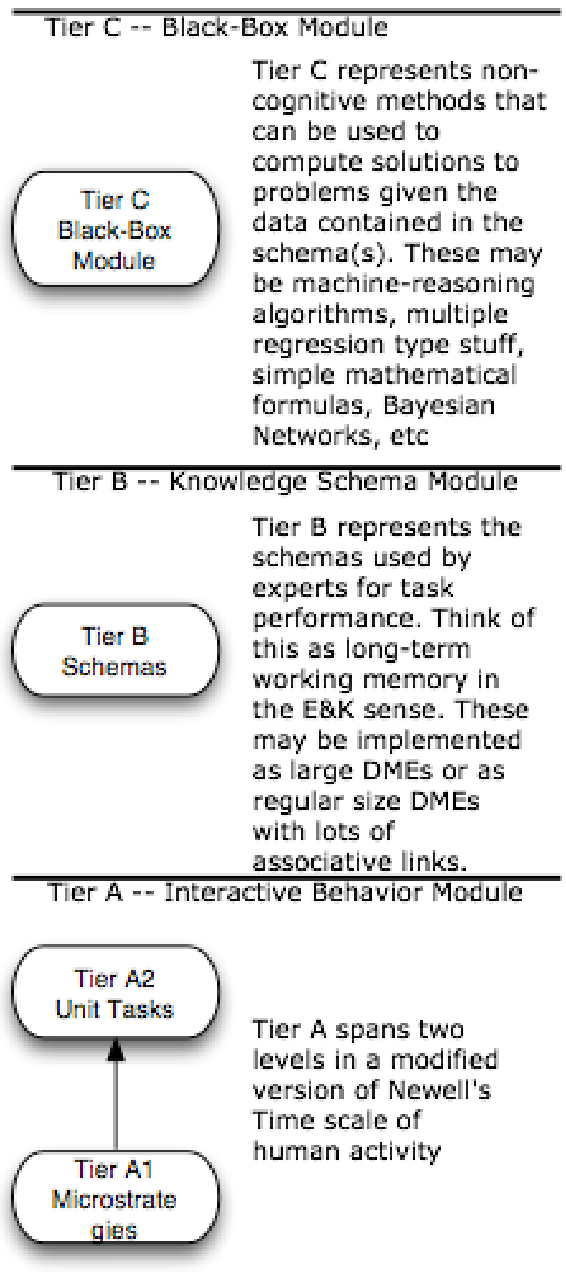
A three-tier architecture for interactive system design emerges from this vision. The *interactive behavior* tier interacts directly with the task environment much in the manner that people interact with it. This is the level of highest cognitive fidelity and represents an off-the-shelf use of ACT-R 5.0 (Anderson, Bothell, Byrne, & Lebiere, 2002). These models adopt the hard and soft constraints approach that is the focus of our basic research efforts. This tier has two subtiers. Tier A1 – is the microstrategy tier. This is the *hard constraints level*. There are only so many ways that a given interactive device can interact with an interactive object. These "ways" are typically constrained by the toolbox of a given operating system. (See, Gray & Boehm-Davis, 2000 for an elaboration of this issue.) Tier A1 is the level of the greatest reusability of productions; that is, the core productions for interacting with interactive objects should be common across models and, as an article of faith, should be reused.

Tier A2 is the unit task level. Unit tasks are more task-specific than are the microstrategies. There may be multiple methods for completing any given unit task – this is the *soft constraints level* where least-cost considerations influence the method chosen to implement a unit task.

Like the humans, the Interactive Behavior Module will interact differently with different scenarios in the same task environment. It will also interact differently accomplishing the same scenario but with different task environments (i.e., different configurations of tools). Having a family of cognitively realistic models of interactive behavior will enable us to quickly predict the influence that various “designed environments” have on human performance.

Success at designing better task environments requires more than models that interact with their environment the way that people interact. The models’ activities must be directed towards the same information acquisition goals as people’s and the information collected must be used for reasoning, problem-solving, or hypothesis generation and testing much as the human. These requirements strain the limits of current computational cognitive modeling. Although modern cognitive science understands much about expert reasoning, complete and cognitively plausible models of expert reasoning are research projects unto themselves. At present, developing the families of high-fidelity cognitive models required to account for the range of human reasoning across multiple task environments would be a separate and daunting line of research.

To finesse this problem, tier B consists of *knowledge schemas* that guide interactive behavior using sets of scenario-specific schemas. These schemas are derived by knowledge engineering techniques (primarily protocol analysis) from human experts as they solve a given scenario. The schemas will guide information acquisition activities and cache the results. (Developing these scenario-specific schemas is, in itself, a major knowledge engineering effort.) As used here, the schema framework has its origins in the Project Nemo work (Gray, 2003).



A three tier architecture for building simBorgs

Schemas may be considered the knowledge

structures that Ericsson and Kintsch (1995) refer to as long-term working memory. There may well be multiple schemas for the same task. (Note that it is probably more accurate to think of sets of schemas rather than one large schema per formal method, but we will talk as if there is a one-to-one correspondence between the methods in the black-box reasoning module and the schemas.) Different schemas may enable the use of different black-box methods.

The third tier, Tier C, is the black-box module. This module may contain a variety of methods for problem solving or reasoning in the task domain. Depending on the domain, the method chosen may be a machine-reasoning algorithm, Bayesian network, multiple regression model, or algebra equations that simply compute an answer. (Indeed, we can envision writing simple arithmetic Lisp functions that compute different decision-making algorithms; i.e., compensatory versus non-compensatory ones, etc.) The essence is that the black-box method provides an answer that a human might give, but the process by which the answer is derived bears no resemblance to human information processing.

The black-box method may be selected based on PG-C considerations. This arrangement would have implications for how we design the black-box module. We may want the black-box method to be fired by Lisp code on the RHS of a production rule. The arrangement we use needs to facilitate two things. First, it needs to facilitate the assignment of credit to the black-box method. Credit determines the cost and probability of success. Second, different black-box methods may be considered at different phases of the data gathering. Rather than having 3 methods that compete at the same time, it may be the case that Method 1 is enabled with less data than the other two methods. In this case Method 1 might compete with a unit task that gathers more data. Method 2 becomes enabled with more data (or different data), it then competes with both Method 1 and a unit task that gathers more (or different) data. In this scheme of things we may want to hardwire an estimate of C for each method and assign each method a given P based on the effectiveness of the method.

Unlike human decision-makers, the black-box methods will always work perfectly. The caveat is GIGO (garbage-in garbage-out). If the first two tiers feed the black-box method old information or wrong information, the black-box method will accurately compute the best answer that can be obtained from the information provided. Hence, this approach promises to provide a general way of exploring how differences in interface design lead people to tradeoff effectiveness versus efficiency in the strategies they use. The goal is to be able to separate the low-level information acquisition and interaction processes from the higher-level, problem-specific expertise required to do a particular job.

The need to support the black-box module will require us to extend ACT-R by the addition of a *black box buffer*. In operation, this buffer will be similar to the other buffers of ACT-R 5.0 (e.g., motor, visual attention, and memory retrieval buffers). Indeed, to the extent that the established ACT-R buffers are grounded in neurocognitive data (Anderson et al., 2002), the addition of a buffer enabling the sending and receiving of information to a black-box module can be conceptualized as a simulated brain implant.

From ACT-R to simBorgs: Cognitive Science to Cognitive Engineering

The idea of building simBorgs is a bold and challenging extension of the current state of the art in computational cognitive modeling. SimBorgs represent a marriage of artificial intelligence algorithms to a high-fidelity embodied cognition capable of interacting with a variety of interface designs and making the same least-effort tradeoffs as human users.

Acknowledgements

Support for the writing of this paper was provided by grants from AFOSR #F49620-03-1-0143 and ONR# N000140310046.

References

Anderson, J. R., Bothell, D., Byrne, M. D., & Lebiere, C. (2002). *An integrated theory of the mind*. Retrieved October 17, 2002, from http://act-r.psy.cmu.edu/papers/403/IntegratedTheory.pdf

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102*(2), 211-245.

Gray, W. D. (2003). *Final Report for Project Nemo: Subgoaling and subschemas for submariners*. Troy, NY: Rensselaer Polytechnic Institute.

Gray, W. D., & Boehm-Davis, D. A. (2000). Milliseconds Matter: An introduction to microstrategies and to their use in describing and predicting interactive behavior. *Journal of Experimental Psychology: Applied, 6*(4), 322-335.