Chapter 10 – COMPLEXITY, HIERARCHY, MODULARITY AND VALIDITY IN HBR ARCHITECTURES

Session Chair: Dr. Kevin Gluck
Warfighter Training Research Division
6030 South Kent Street
Mesa, AZ 85212-6061
USA

Presenters:

Mr. Randy Jones
Soar Technology, Inc.
3600 Green Court, Suite 600
Ann Arbor, MI 48105
USA

Dr. Emiel Ubink
TNO Human Factors
Kampweg 5, PO Box 23
3769 ZG Soesterberg
THE NETHERLANDS

Dr. Rob West
Carleton University
Department of Psychology
1125 Colonel By Drive
Ottawa, Ontario K1S 5B6
CANADA

10.1 INTRODUCTION

In Chapter 8 of *The Sciences of the Artificial*, Herb Simon (1969) described the role of hierarchy in complex systems. Four key points include:

1) Complex systems are nearly always hierarchic;
2) Hierarchy facilitates evolution;
3) Near decomposability is an important characteristic of the hierarchical components of complex systems; and
4) Nearly decomposable, hierarchic structures facilitate comprehension.

In this session we considered HBR architectures from the perspective of nearly decomposable, hierarchical, evolutionary, complex systems.

10.2 SPEAKER PRESENTATIONS

Randy Jones, SOAR Technology Ltd.

Architectures and frameworks have higher level hypotheses and theories that govern them but there is a lot of devil in the detail. Modelling is often much more complex than frameworks suggest. To complicate matters, different architectures have different names for quite similar concepts, processes, and mechanisms. None of the architectures have all the mid-level constructs but they all have several similar key elements. Modelled tasks all have a somewhat similar knowledge structure across architectures. A shared meaning of terms and a common vocabulary would be useful in the community.

No architecture supports all basic representations and processes needed for all HBR needs. Randy expressed the need for a comparative framework to compare functions and features of all different architectures. To highlight the challenge, he noted that in Pew and Mavor (1998), major cognitive architectures are described in the words of the developers who use different terminology for fundamentally similar architectural components. Therefore, a deep understanding of each architecture is required to know which terms and functions are synonyms and which are not.

Randy described his interests and experiences in identifying similarities in processes and representations across cognitive architectures. His initial analysis included BDI, GOMS, and Soar. Randy published this initial work in AI Magazine in 2006. More recently he and colleagues at Soar, in collaboration with Christian Lebiere at Carnegie Mellon, developed an abstraction language called High-Level Symbolic Representation (HLSR) and have been investigating the use of HLSR to simultaneously generate Soar and ACT-R models of the same task. They see HLSR as part of their long-term strategy to create a modular, hierarchical implementation language that is useful for multiple architectures.

Dr. Emiel Ubink, TNO

Modular HBR: The many parts of a human model will in most cases be constructed as a nearly decomposable hierarchy, as described by Herb Simon (see introduction). It seems, however, that most behavioural architectures are an exception to this rule, at least at the functional or behavioural level. In most behavioural architectures several components (blackboard, rule based, etc.) have to be consulted to generate a specific type of behaviour, such as answering the telephone. This shows that on a functional level the behavioural architecture is not at all nearly decomposable. In Emiel's behavioural architecture based on pandemonium theory the components (demons in the pandemonium) are functional units ("behaviour chunks") instead of, for instance, units corresponding to the components in a theory of cognition. The demons in this pandemonium are in competition over a limited set of resources that represent the capabilities of the system and that are needed to generate behaviour. This results in a nearly decomposable system that simplifies the development of complex behaviour, that improves the maintainability, and that allows one to easily plug in or remove specific behaviours.

Variability in Behaviour*:* Variability in behaviour has to do with action selection and action execution (i.e., which actions are taken and how these actions are executed). These "what to do" and "how to do it" questions are often addressed separately. They are, however, very much intertwined, especially with respect to human performance modelling. The answer to the "what" question could very well be influenced by the "how" answer (since I can not run fast enough I will not run at all) and vice versa (if I decide to move to location A, I will have to run). These "what" and "how" questions are so entangled because of the many feedback loops in behaviour, and because both the "what" and the "how" of behaviour depend for a large part on the capabilities of the modelled system, that should be viewed as a limited set of resources. The conclusion is therefore that when modelling human behaviour a way needs to be found to address both questions simultaneously.

A portion of Emiel's presentation focused on the relation of decomposability to validity. A fully decomposable architecture can be validated at the element level. This served as a nice segue into Robert West's presentation.

Dr. Robert West, Carleton University

Robert emphasized the point that evaluations of the validity of models and architectures are required in order to achieve objective, measurable progress in HBR. He and his colleagues at Carleton University are re-implementing ACT-R in Python using modular components that can easily be modified to change the system architecture. This facilitates modification and validation of the separate modules comprising the architecture. Python was chosen to ease code writing instead of using LISP, ACT-R's native language. To make progress, we need to show validity in the approach and the models developed, including predictive validity, but validation is an expensive process that the civilian market seems unwilling to assume. Without validation, we are left with curve-fitting and the best curve-fit solution isn’t necessarily the one that captures the essence of the data best. Architectures, or rather models built in architectures, have to accurately predict human behaviour.

Robert expressed an interest in testing model fit against different datasets and established parameter values. Better ways of fitting model output to empirical data are also needed.

10.3 DISCUSSION

The general discussion focused primarily on modularity, common resources, and validity. Here are some highlights from the points and questions raised.

10.3.1 Modularity

Object oriented programming, service based software architectures, modular thinking (Randy mentioned aspect oriented programming) is the trend in all software development. The need for development of middle level constructs and functionality was raised by both Randy and Emiel.

10.3.2 Common Resources

Architectures for HBR need to be free and open source to promote reuse, particularly by academia, which is well positioned to address the validation question. There is a need to develop open, common agents that do things so that the community can explore them critically.

An on-line resource model-repository would be useful. The NATO RTO SAS 053 Virtual Institute for HBR came up in the discussion, but there was some confusion as to its status. Brian Gore pointed to the need for a place to publish new models. He mentioned a new journal coming out: the International Journal of Human Factors in Modelling and Simulation (Brian Gore is an editor), providing details much as the AMBR project did; NASA also has a book out covering model validation similar to AMBR but in a different context, possibly with a web-based repository also connected to it. Also the journal Cognitive Engineering and Decision Support was mentioned as a place to publish models.

10.3.3 Validity

It can be difficult to differentiate between a good architecture and a good model. How do you validate an architecture for predictive validity? Perhaps construct validity and content validity are sufficient for architectures unless some unique aspect of the architecture forces the model to predict a certain observation. Lots of successful models built in an architecture lends credence to its validity as an architecture, but you have to ensure that the models follow the tenets of the architecture and do not simply develop convenient shortcuts or work-arounds to simplify the model development.

Comparative validity is difficult in the context of architectures and HBR-based models because only the developers really understand the model implementations and approximations, and sometimes even the individual developers don’t have a complete understanding of the entire architecture in sufficient depth. It is hard to compare models of the same phenomena if design decisions are not documented.

Can we validate architectures or can only specific models be validated? An architecture can only be validated through its use in many studies (i.e., validated on a meta-level). Evidence accrues for the validity of an architecture just as evidence accrues for the validity of scientific theories.

Validity, whether at the model level or the architecture/system level is not a binary state. Validation is a process of accrual of evidence and validity is a continuum rather than binary state that is context dependent.

10.4 SESSION RECOMMENDATIONS FOR HBR S&T

We need datasets from a wide and increasing range of tasks in order to objectively establish the broader validity of the cognitive architectures and HBR systems under development today.

Some type of middleware that facilitates cross-HBR comparison is advisable.

The HBR R&D community would likely benefit from a better understanding of productive methods for comparing models across architectures and comparing architectures.

Continuing the trend toward modular HBR systems and cognitive architectures is advisable.

10.5 SESSION RECOMMENDATIONS FOR NATO

There should be some interaction between SAS and HFM to explore potential for collaboration on HBR. The NATO RTO SAS-053 Virtual Institute for HBR, which was previously recommended by RTO SAS-017 panel, is due to wrap up in 2007, but information is required such as the timeline for standing up the Virtual Institute, its envisioned capabilities and suspected limitations. The Virtual Institute for HBR may provide a natural site for locating a repository of models and data that can be made public for modellers in the field.