

A Decomposition Model for the Layered Evaluation of Interactive Adaptive Systems

Alexandros Paramythis¹, and Stephan Weibelzahl²

¹ Johannes Kepler University,
Institute for Information Processing and Microprocessor Technology (FIM)
Altenbergerstraße 69, A-4040 Linz, Austria
alpar@fim.uni-linz.ac.at

² National College of Ireland, School of Informatics, Mayor St.,
IFCS, Dublin 1, Ireland
sweibelzahl@ncirl.ie

Abstract. A promising approach towards evaluating adaptive systems is to decompose the adaptation process and evaluate the system in a “piece-wise” manner. This paper presents a decomposition model that integrates two previous proposals. The main “stages” identified are: (a) collection of input data, (b) interpretation of the collected data, (c) modeling of the current state of the “world”, (d) deciding upon adaptation, and (e) applying adaptation.

1 Introduction

The evaluation of interactive adaptive systems (IAS) is currently receiving considerable attention. This can be attributed, at least in part, to the largely unsolved problems involved in the evaluation of IAS (see, e.g., [2], [4]): There are currently only few empirical studies that evaluate interactive adaptive systems; furthermore, most of the existing studies provide ambiguous results. Although these two claims have been stated repeatedly in the past (e.g., [2]), they are still valid and unresolved.

Recently, there have been attempts at tackling the problem of evaluating IAS by “decomposing” adaptation and evaluating it in a “piece-wise” manner. The premise behind these attempts is that the evaluation of adaptive systems should not treat adaptation as a singular, opaque process; rather, adaptation should be “broken down” into its constituents, and each of these constituents should be evaluated separately where necessary and feasible. The constituents into which adaptation is decomposed are typically termed “layers” and the resulting approach “layered evaluation”.

This paper reports on work-in-progress aimed at combining and expanding upon two of the “layered evaluation” frameworks reported in the literature. The process-based framework presented in [9] discerns four layers that refer to the information processing steps within the adaptation process. The framework has a very clear focus on the empirical evaluation of IAS and has been applied in practice to different adaptive learning courses, including several studies with thousands of users [9]. The sec-

This is a pre-print of the final version which is copyrighted by © Springer Verlag 2005: Paramythis, A. & Weibelzahl, S. (in press). A Decomposition Model for the Layered Evaluation of Interactive Adaptive Systems. In: *Proceedings of the 10th International Conference on User Modeling*. Berlin: Springer

ond framework presented in [7] addresses the issue of formative vs. summative evaluation and, overall, adopts a more “engineering” perspective in the identification of layers, focusing on the different components involved in the adaptation process.

2 A New Model for “Decomposing” Adaptation

Our efforts towards a merging or unification of the two alternative propositions centered on the introduction of a model of decomposition, which: has the widest possible applicability on existing and forthcoming IAS; makes few (if any) assumptions about implementation and architectural properties of the system; but, at the same time, offers a concrete enough guide to evaluation activities.

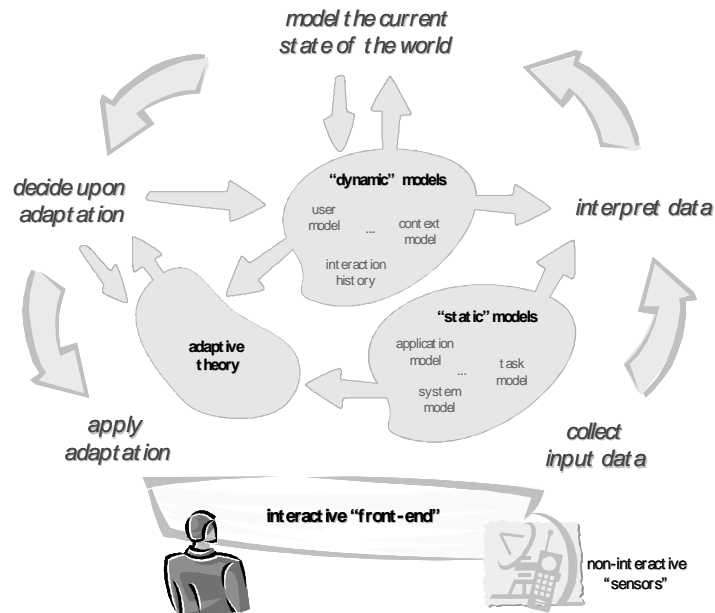


Fig. 1. The proposed adaptation decomposition model.

To arrive at the desired decomposition model, we have examined the common properties of existing models and architecture for adaptation (e.g., [3], [5], [6], [8]). In doing so, we have restricted ourselves to the process-oriented models, so as to allow for the maximum possible degree of flexibility in terms of implementing adaptation (where, in fact, approaches proliferate).

The proposed model is depicted in Figure 1. The main “stages” of adaptation identified are: (a) collection of input data, (b) interpretation of the collected data, (c) modeling of the current state of the “world”, (d) deciding upon adaptation, and (e) applying adaptation (i.e., effecting adaptation decisions). It is argued that each of these adaptation stages needs to be evaluated explicitly, although not all stages can be “iso-

lated” and evaluated separately in all systems. Furthermore, the nature of the IAS will necessarily dictate the relevance of each of these stages.

Note that this figure contains several elements, “internal” to the IAS, which are not part of the model itself, and are included solely to facilitate understanding of the model and support discussion. Further, the decomposition model explicitly makes no assumptions regarding specific approaches to intelligence, or decision-making.

The models potentially maintained by the IAS are separated into two broad categories. The first groups together the IAS’s “static” models (e.g., the system model, the task model, the application model, etc.), which are used, implicitly or explicitly, when interpreting input data. The second category groups together the IAS’s “dynamic” models (e.g., the user model, the context model, etc.), which are updated by the IAS, on the basis of new knowledge that it derives from the interpretation of the input data. These are, typically, the main determinant for adaptation decisions. For both categories, arrows are used to denote potential flows of information. Figure 1 also introduces an entity termed “adaptive theory”. This term, borrowed from [8], is used to refer to the theory that underlies adaptations in the system; the word theory is used here in an informal sense, to represent the totality of a system’s adaptation goals / objectives.

In the rest of this section we present briefly each of the stages that appear in the model and discuss why they need to be evaluated (in isolation or combination) and with what objectives. Due to space limitations, other important aspects of the reported work, such as the introduction of specific evaluation criteria that can serve as “guides” for their respective evaluation layers, are not discussed.

Collection of input data: The “input” data that an interactive system collects is predominantly derived from the user’s interaction with it (another source may be “sensors” not directly or explicitly manipulated by the user). Input data of this nature does not necessarily carry any semantic information. It is in the next stage, and with the assistance of (implicit or explicit) application- and task- models that this low-level data will acquire “meaning”. With respect to evaluation, this category of data is subject only to “technical” assessments which would determine whether factors such as reliability, validity, accuracy, precision, latency, sampling rate, etc. are appropriate for the system at hand. Given the assumption that “raw” input data does not carry semantic value by itself, such assessments may be all that is necessary at this level.

Interpretation of the collected data: This is the very stage at which input data acquire “meaning” of relevance to the system. The distinction between this stage and the collection of the input data is intended to identify and conceptually dissociate the two stages, thus making it possible to address them in isolation. The interpretation process may be straightforward in those cases that there exists a direct, one-to-one mapping between the raw input data and their semantically meaningful counterparts. When the interpretation is unambiguous, and independently of whether it employs any of the system’s “static” models, it can be assessed objectively and in a user-independent manner. Potential problems arise when: (a) the interpretation makes use of assumptions, or (b) the interpretation requires some level of inference. Assumptions and inferences are quite commonly employed in existing IAS, mainly due to the lack of additional data that can better describe the context of interaction.

Modeling of the current state of the “world”: This stage concerns the derivation of new knowledge about the user, the user’s group, the interaction context, etc., as well as the subsequent introduction of that knowledge in the “dynamic” models of the IAS.

There is a definite overlap between this stage and the interpretation of the input data; in fact, in several cases, there is no “second-level inference” in adaptive systems, which simply go from interpreting the input data to representing those interpretations in an appropriate model. The main evaluation goal for this stage is validity of the interpretations / inferences. Secondary yet important concerns that also relate to the modeling process include: (a) comprehensiveness of the model; (b) redundancy of the model; (c) precision of the model; and, (d) sensitivity of the modeling process.

Deciding upon adaptation: During this adaptation stage (also referred to as the “eference” stage), the IAS decides upon the necessity of, as well as the required type of, adaptations, given a particular “state” (as expressed in the various models maintained by the system, or directly from input data). We make a very clear distinction between this stage and the next (see “Applying adaptation decisions” below), as a way of facilitating the conceptualisation of the steps that are involved in the derivation and application of adaptation decisions. A “rule of thumb” we propose for separating the two stages is: decisions made at this stage are mainly at the semantic and upper syntactic level of the interaction results; further decisions made while effecting adaptation belong to the lower syntactic, or to the lexical / physical level of interaction. The goal in making this distinction is to foster the separation of the adaptation theory from decisions that represent a typical interaction design task, rather than a particular adaptation artefact. The primary aim of this evaluation step is to determine whether the adaptation decisions made are the optimal (e.g., necessary, appropriate, subjectively accepted by the user) ones, given that the user’s properties have been inferred correctly.

Applying adaptation decisions: This stage refers to the actual introduction of adaptations in the user-system interaction, on the basis of the related decisions. Although typically subsumed by adaptation decision making in the literature, this stage may be varied independently of the decision making process, e.g., to account for different adaptation strategies. More importantly, this stage usually “hides” a level of adaptation (i.e., the transformation of possibly high-level adaptation decisions to a “concrete” form experienced by the user), which only too often, and in several cases mistakenly in the authors’ opinion, gets evaluated in tandem with the higher-level decision making stage. The evaluation criteria that are applicable at this stage depend very much on the type of adaptation effected. In most cases, traditional evaluation criteria, such as usability, will be highly relevant. The identification of these criteria can only be performed on a case-by-case basis, although the application of some general criteria (e.g., timelines, obtrusiveness, level of user control) may be feasible.

Evaluating adaptation as a whole: The “piece-wise” evaluation of adaptation, as proposed herein, can provide valuable insight into the individual adaptation stages through which an IAS goes. However, what is still missing is the “big picture” – the evaluation of the primary adaptation theory (or theories). To assert whether such high-level theories (or, seen from a different perspective, hypotheses) hold true, one needs metrics that transcend the layered evaluation of adaptation, as this has been discussed so far. Such metrics must adequately capture the application- and adaptation- domains, to be able to more holistically assess the “success of adaptation”. Browne, Norman and Riches [1] have proposed that this problem be addressed by: (a) articulating and assessing against the system’s objectives, and / or (b) assessing indirectly against the underlying theory. According to these authors, many of the objectives of an adaptive system can be expressed as lists of purposes, which, in turn, can be

loosely interpreted as the collection of “reasons” that led to the introduction of adaptation in the system. Metrics and assessment methods can then be devised to measure the extent to which the stated objectives are met.

3 Summary

The postulation of layered evaluation of IAS is that adaptation needs to be decomposed and assessed in layers in order to be evaluated effectively. The decomposition model proposed here takes a process-oriented approach to the decomposition, identifying the logical stages through which adaptation progresses. A brief rationalization of the decomposition and a preliminary set of criteria have also been put forward.

An important point we would like to make about the proposed decomposition is that it is neither the only one feasible, nor, necessarily, the most appropriate one for all types of assessment of IAS one might want to perform. For instance, it would be possible to decompose adaptation on the basis of the software components involved in its implementation. The same is true for the level of granularity to be employed.

References

1. Browne, D., Norman, M., Riches, D.: Why Build Adaptive Systems? In: D. Browne, P. Totterdell, & M. Norman (Eds.): *Adaptive User Interfaces*, London, Academic Press (1990) 15-58
2. Chin, D. N.: Empirical evaluation of user models and user-adapted systems. *User Modeling and User-Adapted Interaction* (2001) 11(1-2), 181–194
3. de Bra, P., Houben, G.-J., Wu, H.: AHAM: A Dexter-based reference model for adaptive hypermedia. In: Tochtermann, K., Westbomke, J., Wiil, U. K., Leggett, J. (Eds.), *Proceedings of the ACM Conference on Hypertext and Hypermedia* (1999) Darmstadt, 147–156
4. Höök, K.: Steps to take before intelligent user interfaces become real. *Interacting with Computers* (2000) 12(4), 409-426
5. Koch, N., Wirsing, M.: The Munich Reference Model for adaptive hypermedia applications. In de Bra, P., Brusilovsky, P., Conejo, R. (Eds.), *Proceedings of the Second International Conference on Adaptive Hypermedia and Adaptive Web Based Systems*, Málaga, Spain, AH2002 (2002) Berlin, Springer, 213–222
6. Oppermann, R.: Adaptively supported adaptability. *International Journal of Human Computer Studies* (1994) 40(3), 455–472
7. Paramythis, A., Totter, A., Stephanidis, C.: A Modular Approach to the Evaluation of Adaptive User Interfaces. In: S. Weibelzahl, D. Chin and G. Weber (Eds.) *Empirical Evaluation of Adaptive Systems*. Proceedings of workshop held at the Eighth International Conference on User Modeling in Sonthofen, Germany, July 13th (2001) Freiburg, Pedagogical University of Freiburg, 9-24
8. Totterdell, P., Rautenbach, P.: Adaptation as a Problem of Design. In: D. Browne, P. Totterdell, M. Norman (Eds.), *Adaptive User Interfaces* (1990) London, Academic Press, 61-84
9. Weibelzahl, S., Weber, G.: Evaluating the inference mechanism of adaptive learning systems. In: *User Modeling 2003: Proceedings of the Eighth International Conference, UM2003* (2003) Berlin, Springer