Towards Improved User-Product Testing With Cognitively Enhanced Scenarios

Wilhelm Frederik van der Vegte  
Delft University of Technology  
Delft, The Netherlands

Niels C.C.M. Moes  
Delft University of Technology  
Delft, The Netherlands

ABSTRACT

With the increasing information-intensiveness of products, users are challenged with expanding options and possible ways to interact. Rapidly escalating numbers of possible user-operation sequences hinder designers in anticipating all possible (unacceptable) outcomes. Interactively simulating product models with human subjects to explore all options is not practicable. Virtual simulation with computer models of users can open the way towards faster-than-real-time performance and investigation of massive numbers of interaction sequences. This paper reports on opportunities to improve realism of virtual-use simulations by incorporating knowledge about the workings of the human brain. We elaborate how, in particular, cognitive-architecture simulations developed by cognitive scientists and error phenotypes identified in human reliability analysis (HRA) can extend a virtual-use simulation approach that we have proposed in foregoing work, by offering the perspective of generating interaction sequences with erroneous user actions unforeseen by the designer. We outline how such an integrated system can be implemented and also discuss validation issues.

1 INTRODUCTION

The rise of smart systems and ubiquitous computing has led to increasing information-intensiveness of products. As a consequence, issues that used to be associated with complex systems such as nuclear plants and airplanes are more and more becoming issues of everyday products such as cars, personal computers and even electric irons or microwave ovens. One of these issues is unforeseen use. The multitude of options that information-intensive products present to users can invoke erroneous user actions governed by cognitive phenomena such as distraction, forgetfulness, information overload and learnability [1-4].

The growing complexity of possible interaction patterns can rapidly escalate the number of possible variations of user-operation sequences, and for designers it becomes increasingly difficult to foresee all possible outcomes, which might include unacceptable performance, failure, and even fatalities. To help solving this problem, we aim to develop a method to predict possible outcomes of use during design, and thus allow anticipation before the product/system is realized. In order to deal with the large numbers of variations we have focused our efforts on computer simulations.

In previous work [5-7] we showed how complex interaction processes can be virtually tested by controlling conventional engineering simulations with use scenarios. Virtual testing allows early design evaluation since it can be done completely on a computer without deploying human subjects or physical prototypes. A scenario describes ‘a way to use a product’. By unifying multiple ways of use in a timed hybrid automata (THA) representation, scenario bundles address the fact that most products are used in multiple ways, and also multiple times. In recent work we have reported how, specifically in the case of information-intensive products, simulations much faster than real time can be realized using a unified THA-based representation for all processes involved in using a product, including processes governed by the laws of physics [8]. The key advantage of simulations much faster than real time is that they allow large-scale exploration of the permutational variety of possible options.

However, a weakness of our approach in its present form is lack of realism of the user behaviors in the simulation. Currently user actions are ‘programmed’ into the scenario bundle (for instance, by taking the intended user manual as a starting point) based on conjecture by the designer, who can specify some additional variation by randomizing certain parameters. What we would like to simulate as well, however, is the effects of human behaviors beyond what the designer can think of—in particular, human ‘errors’ such as forgetting to check the status of the product when this is required, or taking too long to complete an action. By making simulations more realistic in this respect, we expect that designers can better...
anticipate and thus prevent user errors and/or their effects. After all, there is general agreement that primary responsibility to prevent human ‘errors’ is not with the end users themselves but with the designers of a system, since a design should inherently suit the capabilities of its users [e.g., 1,3].

Generally user ‘errors’ can be traced back to the workings of the human brain and the central nervous system, and the related human behaviors are studied by scientists in areas such as cognitive psychology, human motor science, human reliability analysis (HRA), and human-computer interaction (HCI). Some of their efforts have led to simulation algorithms to or to formalizable knowledge that can otherwise be used in simulations. Most notably, cognitive architectures (CAs) such as ACT-R [9], EPIC [10], and SOAR [11], which have been developed by cognitive scientists, are interesting because they have been empirically validated and allow simulation of some of the key phenomena behind human error, such as learnability multitasking, learning, attention and distraction [12-14]. CAs are production-rule based blueprints of the operational structure of cognition. They predict the time the human brain needs for processing perceptual input and for directing motor operations. However they cannot predict the reasoning that determines which operation is performed in which situation, and this is why we still need scenario bundles. CA-based simulations start from a given task decomposition, i.e., an idealized sequence of operations as it can be extracted from an instruction manual.

In this paper we explore the opportunities to utilize CAs and related knowledge collected in associated areas of science with the objective to increase the realism of simulations involving virtual human users – in particular, but not exclusively, in combination with our approach that is briefly described in Section 3.

The paper is structured as follows: In the next section we discuss related achievements by others. Then, in Section 3 we briefly describe our approach of scenario-bundle based simulation (SBBS) without the proposed cognitive enhancements. Section 4 discusses the key cognitive aspects that have to be considered. In Section 5 we elaborate on our proposition to extend SBBS with cognitive abilities, and in Section 6 we wrap up with discussion and conclusions.

2 RELATED WORK

There have been few other efforts to operationalize CAs in simulations of the use of products. Since CAs originate from cooperation between computer scientists and psychologists [15], their application has long been limited to HCI [16]. With the emergence of ubiquitous technologies, everyday products are becoming increasingly information-intensive [17] and their use processes involve behaviors and interactions that are uncommon in HCI and have not been included in CAs. CAs offer effective shortcuts to simulate interactions with keyboards, monitors, etc. with disregard of physics [10]. Recently, similar specific solutions have been developed for driving cars [e.g., 18,19], and piloting aircraft [20] but so far no approach seems to exist that is widely applicable to information-intensive products and that has been devised for use in a design context. Possibly the only exception is the ACT-R/DHM project at Mississippi State University, where researchers have attempted to enhance a 3D digital human model (DHM) with cognitive modules from ACT-R [21,22]. However, there are no reports on this project since 2009. At that time, the researchers had only considered modules for visually locating interaction elements, and had not considered further stages of human-product interaction.

3 SIMULATING USE PROCESSES WITH TIMED HYBRID AUTOMATA: SBBS

In [8] we presented a scenario-bundle based simulation (SBBS) approach to simulate the use of products based on a unified THA-based representation of actions and processes performed by – and taking place in – humans and artifacts. The idea behind this approach is to enable simulations much faster than real time by replacing the nowadays popular 3D product-simulation models for physics simulation with simple THA representations. 3D Simulation approaches, such as the ones based on which finite-element (FE) and multibody systems, are usually complex due to their dedication to accuracy, geometric completeness and, in the case of FE and other approaches based on spatial discretization, (approximation of) continuum physics, which requires keeping track of many points in space at many points in time. These simulations allow designers to investigate the dynamics of spatial changes in humans and products, and spatial manipulations that designers are supposed to perform. However, in the case of information-intensive products, where successful use depends on information exchange and processing rather than on physics or human dexterity, such investigations are often not needed.

The logic of information exchange and processing can be described as discrete processes using THA. Advanced automata representations such as statecharts [23] can also be used to describe continuous behaviors based on (systems of) equations [e.g., 24]. Actually, the addition of continuous behaviors is what makes the automaton a hybrid one (i.e., the H in THA). This allows us to describe processes governed by physics – that is, up to a certain level of complexity, as long as keeping track of many points in space at many points in time is not needed.

As a consequence, we can simulate use processes based on...
models and specifications\(^1\) that are solely based on THA and that cover both the user and the product (as well as other involved artifacts and humans). As an example, Fig. 1 shows the models and specifications that we created with Matlab Simulink Stateflow to simulate the use of an espresso machine. They are uniformly represented as statecharts that are executed in parallel. Figure 2 shows an excerpt of the scenario bundle that describes actions by the virtual human user, which have been specified as reactions to supposedly perceived feedback from the espresso machine (e.g., a lamp blinks, the sound of the pump can be heard, etc.). The actions are represented as states (rounded rectangles) containing control commands and the perceived feedback is represented as transitions (arrows) to other actions, which take place based on specified conditions.

By executing this statechart together with similar models and specifications that describe the behavior of the espresso machine, simulations can be run at speeds up to \(5,000 \times\) real-time, so that, for instance, a year of using the product can be simulated in less than two hours. An example of possible simulation output is shown in Fig. 3. Based on such output, the designer can, for instance, optimize the parameters related to the power-save mode in relation to targeted energy savings and assumed break-taking habits of the user. In the absence of functionality to predict durations of human actions, we built in custom functions to randomize durations within a given range, to be estimated by the designer. Similar functions are included to allow randomization of choices between different operation modes (e.g., lungo/large cup or espresso/small cup) and, for instance, filling level of the reservoir.

4 CONSIDERATION OF COGNITIVE CAPABILITIES IN USE-PROCESS SIMULATION

By means of scenario bundles, SBBS offers very limited support of cognitive activity related to human-product interaction. In that respect, a fully capable simulation approach should include the following three main aspects of cognition:

- **Decision-making:** Scenario bundles describe high-level decision-making but they disregard lower-level brain activity related to prioritizing and evaluating production

\(^1\) We distinguish physics models and logical specifications because physics models are typically based on validated relationships or derived from laws of nature, while logic is typically described based on assumptions or man-made program code.
rules, accessing short-term/long-term memory, dividing attention, etc. Models of these activities belong to the core modules of CAs. Decision-making and the related cognitive processing are the source of the human ‘errors’ that we aim to reveal in simulations. In HRA, cognitive errors are considered at two levels: phenotype and genotype [25]. Error phenotypes (e.g., Fig. 4) can be observed in overt action, i.e., they correspond to aberrations at scenario level. However currently, a scenario bundle can only produce aberrations foreseen by its author. Genotype corresponds to the cognitive mechanisms causing errors and thus to CA-level processing. HRA and CA researchers [1,25-33], have presented genotype mechanisms and models, which are however not (yet) universally applicable in cognitive simulations.

- Motor control: in [7], we provisionally used unvalidated algorithms to simulate posture-to-posture transitions. Meanwhile, others introduced validated algorithms [34-36]. These algorithms however exceed the functional capabilities that we currently need, since they offer full control of 3D spatial models. To extend capabilities of the HCI-centered motor models in CAs without considering full 3D control, Fitts’s law [37] has been suggested [38]. In addition, recent refinements of Fitts’s law and other invariants of motor control have been proposed [39-41], which can also be used to predict durations.

- Perception: regarding the cognitive aspects of perception, it is crucial to predict whether relevant events are detected, and if so, how much time is needed. In CAs, perception has been simplified by discretizing colors, zones of vision, etc. [9,10].

5 EXTENDING SBBS WITH COGNITIVE ABILITIES

To extend scenario-bundle based simulations with cognitive abilities, we propose to link them to a cognitive architecture that intercepts the inputs to and outputs from the scenario bundle, and, based on computation of durations of cognitive processes, imposes delays (i) before the scenario bundle can process its inputs from the artifact model and (ii) before it outputs the results of its processing to the artifact model. From the available CAs we propose to use ACT-R, since this architecture in particular has built-in perceptual and motor mechanisms that allow its models to interact with external simulations [18].

The architecture of ACT-R consists of four main modules, which are connected to a central production system through their respective buffers; most of these entities have been associated with particular regions of the brain (Fig. 5). The intentional or goal module keeps track of current goals and intentions, the declarative module retrieves information from memory, the visual module identifies objects in the visual field, and the motor module controls the limbs. The central production system coordinates in the behaviors of these modules. It can only respond to a limited amount of information that is deposited in the respective buffers of the modules by request of the central production system [9]. In ACT-R, user actions from a scenario appear as tasks, and scenario bundles combine these as structures of goals [cf. 42]. During simulation, the scenario proceeds from state to state based on fulfillment of conditions (user-perceived events). Thus, towards ACT-R the scenario acts as a new manifestation of the intentional module, feeding the goal buffer with upcoming tasks. Based on these, production rules control the motor module, which computes times for performing each successive task.

Novel extensions of ACT-R in the context of linking to SBBS concentrate on the motor module, adding currently missing models for operating everyday products based on motion invariants. The new models must (i) predict motor-task durations and (ii) prepare outputs compatible with simulations.
of products and use environments. As is common practice for models that are newly added to CAs, they will have to be validated with human subjects.

From the error phenotypes in Fig. 4, the above combination of ACT-R and scenarios can only predict delays and, possibly, premature actions. One way to include more phenotypes is to (further) develop genotype models for ACT-R’s core modules based on [1,25-33]. Instead, we propose to explore and generate error phenotypes from scenarios. Manually assisted preparation of error phenotypes from a taxonomy by Reason [1] for ACT-R’s inputs has been described in [43]. Hollnagel’s taxonomy (Fig. 4) has the advantage over other taxonomies (e.g., 1,44] that it refers to the actions readily appearing in scenarios, which provides the opportunity to develop a novel module that generates errors automatically, producing deviations from the regular scenario by systematic reordering and reconnection. We propose to elaborate on this idea, aware of the following considerations:

- One group in Fig. 4 defines actions not regularly included. Additional actions foreseen in the scenario bundle can be explored but further meaningful actions will be hard to generate algorithmically.
- Exhaustive exploration of possible errors causes combinatorial explosion of the number of scenarios. Therefore, exploration should automatically stop at a given stage.
- We cannot directly assess probabilities of generated scenarios, which would be possible with validated genotypes. However, we aim to derive rules to predict probabilities of phenotypes from experiments involving human subjects.

Figure 6 gives an overview of the intended system. To connect its components in an initial prototype setup, we will use the simulation environment Matlab Simulink, which has proven capabilities to link CAs to statecharts (in Stateflow)

We expect that the error phenotype generator for alternative scenarios can be programmed using the Stateflow API.

6 DISCUSSION AND CONCLUSIONS

Above we have sketched a first outline of how we think that integrated simulation of use processes can be enhanced with cognitive capabilities. Realizing a system as we have proposed above is obviously a complex endeavor that requires further investigation. Therefore, we think that our first efforts should be aimed at validating the two basic issues: (i) is it feasible to identify a manageable collection of typical interactions that cover the use of most everyday products (outside HCI), and can the motor control related to these interactions be modeled as extensions of ACT-R’s motor modules, and (ii) can a setup as is shown in Figure 6 predict the occurrence of real-life human errors? An interesting third question is, whether it is possible to derive rules concerning probability of typical error phenotypes.

We realize that even in the long run, validation for arbitrary information-intensive products is not a realistic objective. It probably means that if we can make the system work, it will have to prove itself in design practice. Another, related challenge that we have to deal with in future work is the evaluation of the effectiveness of the approach.

In the above proposition we have focused on realizing fast simulations by largely ignoring 3D geometry and using simplified models of physics processes. We expect that such simulations will help designers to identify and anticipate possible human errors in the early stage of designing information-intensive products. Another interesting line of research would be to explore the possibilities of connecting similar cognitive simulation capabilities to 3D human models and artifact models. This may provide interesting insights in application areas where human dexterity is important, such as in sports or surgery.

REFERENCES


[37] Fitts, P.M., 1951, Human engineering for an effective air navigation and traffic control system, National Research Council, Washington D.C.


