IMET - International Conference on Interactive Media, Smart Systems and Emerging Technologies (2023)
N. Pelechano, F. Liarokapis, D. Rohmer, and A. Asadipour (Editors)
Short Paper

Recognizing User Behavior from Interactions for Adaptive Consumer Information Systems

S. Lengauer[1], M.A. Bedek[2], C. Kupfer[2], L. Shao[1], D. Albert[2], and T. Schreck[1]

[1]Graz University of Technology, Institute of Computer Graphics and Knowledge Visualization, Austria
[2]University of Graz, Institute of Psychology, Austria

Abstract

Consumer Information Systems, which experience widespread application, benefit substantially from adapting the conveyed information to specific user needs, by addressing various impairments such as color blindness, deficient preknowledge, and/or graph illiteracy. Ideally, to allow for an unperturbed exploration process, the system automatically recognizes and responds to the need for adaptation. While it has been shown that users' interactions with a system can be leveraged to this end, there exists no generalized taxonomy covering all possible interactions/processes and how they relate to each other. This paper garners different interactions, defined in the literature, and classifies them regarding complexity and inter-dependencies in a 'processes landscape'. Using this landscape, we outline a concept how low-level interactions (e.g., 'Clicking', 'Typing') can be combined with context-sensitive ones (e.g., 'Hovering') to estimate high-level behavior such as 'Reading' or 'Exploring'. Knowledge of the latter allows a system to intervene and adapt in a reasonably manner.

CCS Concepts

• ***Information systems*** → ***Personalization;*** • ***Human-centered computing*** → ***Interactive systems and tools;*** *Systems and tools for interaction design;* • ***Applied computing*** → *Health care information systems;*

1. Introduction

Consumer Information Systems (CISs) experience increasing popularity as they are able to replace conventional information transfer in domains such as health, customer service, or E-commerce, with a fraction of the required costs and man-hours. While CISs appear in physical form, e.g., brochures, flyers or operation manuals, the greatest potential in scalability and range lies within web-based systems. Those can convey various types of information (textual, pictorial, multimedial, etc.) and can potentially adapt to users' needs in terms of abstraction level, visual complexity, complexity of information and such; thereby addressing individual impairments such as color blindness, low graph literacy, deficient preknowledge, etc.

The ultimate goal is to provide a CIS that can be used efficiently and autonomously by all kinds of users. One model of how information can be gained by users as efficiently as possible from various sources (patches), such as visualizations or text snippets, is known as *information foraging theory* [PC99]. Efficient information foragers maximize the rate of gaining valuable information within the available time. The available time can be spent on the search for new patches, bearing potentially valuable information (*between-patch processing*), and on the more elaborated processing of these patches to extract relevant information (*within-patch processing*), i.e., by actually reading and interpreting the text snippets or by inspecting and comprehending visualization. When spending the available time solely on the processing of one or only few patches, potentially even more valuable information from other patches will be missed. On the contrary, when spending time only on superficial browsing for new potentially relevant patches, less information can be extracted. Thus, a balanced distribution between these processes is considered as efficient.

A CIS should both adapt to user needs and encourage a balanced exploration process, which can be achieved through subtle interventions such as tutorials, explanations of the functionalities and features, recommendations of tailored content, etc. A prerequisite for this objective is the correct identification of what a user's needs are, since users' digital proficiency and information need are often unknown. While this can be accurately determined by psychologists in a supervised testing, it is much more challenging for an unsupervised web-based exploration process. However, explicit assessments would impair a user's exploration process and potential flow experience [CCAN14]. For an unperturbed exploration process, this knowledge must be obtained purely unobtrusively by analyzing users' behavior [CLWB01; GW09]. In recommender systems, such behavioral patterns are referred to as *implicit interest indicators* [CLWB01]. In the context of technology-enhanced learning and game-based learning, such implicit assessment by interpreting interaction patterns and behavior is called *stealth assessment* [Shu11; SKS22]. Thus, the intermediate goal is to classify a user's state of mind based on interaction patterns – e.g., is he/she behaving confidently or confused – and respond with alternative visualizations/representations accordingly. This enables us to employ multifaceted adaptation principles and to provide tailored interventions if required (such as support offerings, explanations and tutorials, reducing or increasing visual complexity, etc.). In terms of inputs, we have to rely on mouse and keyboard inputs, which can be tracked automatically [ASW06]. As it has been shown that users' mouse movement is strongly correlated with their gaze [Coo06; HWB12], various behavioral patterns have already been described, comprising technical processes such as resting and action [HWB12] or cognitive processes such as reading, examining or exploring [CLWB01; GW09; RF07; YaKSJ07]. Yet, a largely unaddressed problem is how different processes relate to each other and how they could be employed to detect a need for adaptation.

Based on a review of the literature, and informed by the development of a concrete CIS in the health domain[†], we propose a concept to bridge this gap by estimating users' cognitive state. We assume that it can be roughly estimated using their interaction patterns, paired with the exploration history (e.g., alternations between reading, scrolling, changing of representation) and the context within which a pattern occurs. The aim of this paper is to bring processes/events defined in literature into a wider context, unify diverging terminology, and propose additional (context-sensitive) interactions in an encompassing processes landscape. Our work is a first step toward a holistic framework for describing the interaction processes and potentially relevant aspects of it which can be captured and leveraged for adaptation.

2. Related Work

Several publications focus on the relation between gaze and cursor movement [Coo06; GSL*02; HWB12], which appears to have a strong correlation. Hence, the latter is often used as an implicit interest indicator [CLWB01]. E.g., Liu and Chung [LC07] use online mouse movement to estimate student attention while browsing through educational websites. Guo and Agichtein [GA10] propose a combination of search- and user-model, incorporating both context information and basal hardware inputs to infer user intents. Huang et al. [HWB12] investigate the correlation between gaze and cursor alignment in a comprehensive user study. Specifically, they analyse the behavior of users doing Web searches and define interaction patterns indicating 'Inactive', 'Reading', 'Action' and 'Examining' states. Buscher et al. [BvED09] evaluate the applicability of display time as implicit relevance feedback. Using eye-tracking for validation in a user study, they conclude that segment-wise display times (resulting from scrolling actions) pose a valid interest measure. Goecks and Shavlik [GS00] assess the relevance of web pages for users by observing various mouse events which are fed into a neural network. Interactions have also been analyzed to predict the efficiency of a user in a visual search task [BOZ*14] or to infer his/her attention [OGW19]. Ha et al. [HMGO22] propose a benchmark, evaluating different user models defined over low-level interactions.

[†] Novel Concepts and Testbed Systems for Visual Adaptive Consumer Health Information (A+CHIS, https://apchis.cgv.tugraz.at/)

It has been shown that user behavior can be analyzed interactively for recommending reasonable visualization alternatives [GW09; MVT16], dynamic pre-fetching strategies [BCS16], or even to detect fraudulent behavior in filling online forms [WVS*21]. While those concepts are tailored to specific applications, we propose a more generalized methodology, applicable to any CIS.

3. Processes Landscape

We categorize all common interaction processes into three separable tiers with increasing complexity and duration, ranging from basal *hardware* (HW) inputs, over *context-sensitive actions/events* (AE) to *cognitive processes/states* (CS), as illustrated in Fig. 1.

HW Inputs The first tier comprises instantaneous actions, such as click or key press, triggered by the input devices mouse and keyboard. While other input devices such as scanners, drawing pads, webcams, etc. exist, we limit our taxonomy to the aforementioned, as these are the most essential ones for web-based exploration. The actions of the first tier are characterized by the fact that they are completely distinguishable and can be unambiguously tracked. That is, we define the actions 'Clicking', 'Moving', 'Scrolling', and 'Pressing/Releasing' for the mouse and the keyboard, respectively.

Context-sensitive Actions/Events The processes of the second tier base on the inputs of the first tier but require additional information regarding the context in order to detect them. A 'Hover' event, e.g., can not be recognized purely from mouse movement or scrolling, but it needs to be determined additionally that the cursor is over a 'hoverable' element. In contrast to the first tier, the boundaries between processes are fuzzier as it comprises processes of various degrees of complexity and duration, which can also happen simultaneously and build on each other. The entries of this tier consist of both actions already mentioned in literature and actions/events defined by our own, which we deem relevant for determining a cognitive state. In Fig. 1 we propose a rough hierarchy and ordering in terms of complexity. Specifically, we have:

AE.1/2/5 'Fail' events subsume the situation that a user tries to click, scroll or type in a situation where it is not supported by the system. Although not being mentioned by other sources, we believe that they can be an indicator for a 'Lost/Confused state'.

AE.3/4 In terms of keyboard inputs, we differentiate between the two cases 'Shortcut' and 'Typing' (when keyboard input is possible). With the prior, a user tries to invoke another event such as 'Scrolling' [CLWB01], 'Zooming', or 'Searching'.

AE.6–9 Several user interactions result in a change of the interface, such as expanding or collapsing certain elements (**AE.6**), triggering additional information in tooltips (**AE.7**); or a change of the content by means of keyword searches or the application of various filters [GW09] (**AE.9**).

AE.10–12 In terms of cursor patterns (**AE.10**), we differentiate between those resulting in a 'Conscious Action' (**AE.12**) and those exhibiting non-target-oriented behavior (**AE.11**). For the prior, Huang et al. [HWB12] classify 'action' behavior as the situation where a user is "interacting with elements of the page or with the Web page". I.e., things like consciously clicking a link, editing

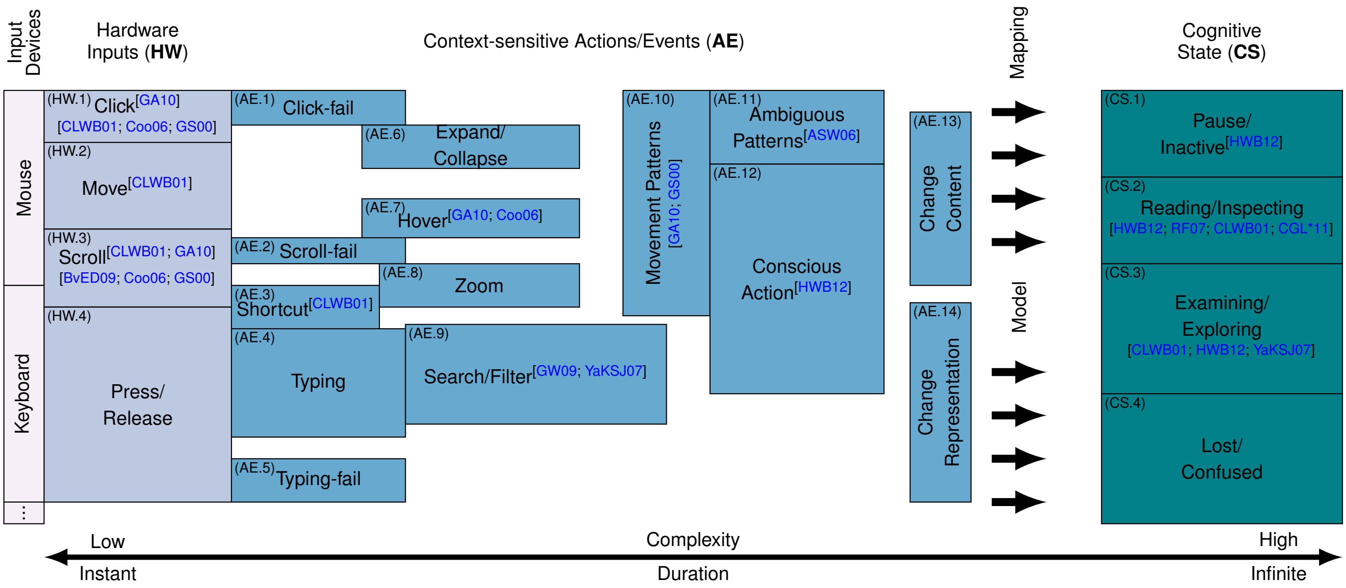

Figure 1: *Our proposed processes landscape introduces a hierarchy into both established processes (see references) and such defined by our own. They are sorted horizontally by complexity/duration, while the vertical dimension indicates possible relations.*

the query in a search box or dragging the scrollbar. As opposed to that, Arroyo et al. [ASW06] state that "slow and arched trajectories as users move their mouse would indicate an ambiguous state of mind", which we dub 'Ambiguous Patterns' (**AE.11**).

AE.13/14 These events subsume all actions/events resulting in a change of the system's displayed content (**AE.13**) or a change of its representation (**AE.14**). I.e., they can be readily defined over the above-mentioned events (e.g., 'Hovering', 'Zooming' or 'Searching/Filtering', respectively).

Cognitive States The third and final tier comprises four cognitive processes – 'Pause/Inactive', 'Reading/Inspecting', 'Examining/Exploring', and 'Lost/Confused' – relating to a user's cognitive state. Although we are aware that the categorization of the faceted and complex human mind in these four generic states is a considerable simplification, we note that this is the level of detail necessary for a CIS to respond with reasonable interventions (Sec. 4). Hence, a 'finer' granularization is neither necessary nor desirable.

CS.1 This is the most easily recognizable state and describes the situation where the cursor remains unmoved for some time and no HW inputs are received whatsoever. Huang et al. [HWB12] define 'Inactive' as the cursor staying still for at least one second.

CS.2 Detecting whether a user is in a 'Reading' state is the subject of several related publications [CLWB01; HWB12; RF07]. All authors agree that user behaviors while reading differ. I.e., Claypool et al. [CLWB01] observed that some users employ the cursor as a reading aid while others use it solely for clicking. Rodden and Fu [RF07] observe three different cursor movement patterns during reading in a user study: (i) *keeping the mouse still while reading*, (ii) *using the mouse as a reading aid*, and (iii) *using the mouse to mark an interesting result*. Huang et al. [HWB12] define a concrete pixel-base rule-set for determining a reading state. Since reading explicitly refers to the consummation of textual content, we propose to extent this category by an equivalent for the consummation of pictorial content, which we dub 'Inspecting'. We suspect it to be strongly correlated with a cursor hovering over the respective element or resting nearby.

CS.3 Huang et al. [HWB12] define this state as the situation where the cursor experiences movement, but the user is neither in a 'Reading' nor an 'Action' mode.

CS.4 Although mostly overlooked in the classification of cognitive states in related publications, user evaluations conducted by our own have shown, that oftentimes users are overwhelmed by the currently displayed information/visualizations. I.e., we believe it is essential to differentiate this state from 'Examining', as it requires a different response from the CIS. We assume the transition between those two states to be smooth, but certain context-sensitive actions/events such as a disproportionate high number of click/scroll/typing-fails or slow/arched cursor patterns, as described by Arroyo et al. [ASW06], could be a strong indication of a confused cognitive state.

To summarize, even though characteristic patterns for some of the cognitive states are provided in literature, we observe that the borders between them are blurry.

4. Discussion

Our concept is to model the cognitive state over both the context-sensitive actions/events and HW inputs. Note that they cannot be modelled directly over the latter, as the intermediate context-sensitive level conveys additional information. I.e., depending on the context, a mouse input can result in, e.g., a 'Hovering' or 'Zooming' – something that is not observable from the raw HW inputs. Hence, we believe that the modeling and capturing of these intermediate level processes is an integral part of the overall concept. The classification of a user's cognitive state into the four distinct cases, 'Pause/Inactive', 'Reading/Inspecting', 'Examining/Exploring', and 'Lost/Confused', is debatable, as other groupings are thinkable [YaKSJ07]. This selection, however, was made with the intended use case of adaptation to a user's need in mind. I.e., the defined states directly map to measures to be taken regarding user support, such as (**CS.2**) no intervening measures to avoid disrupting the user's focus, (**CS.3**) subliminal proposing alternative visualizations/contents, or (**CS.4**) intervening by providing guidelines, tutorials, visualization alternatives, or such. Besides that, a CIS should encourage a proper balance between and **CS.2** and **CS.3**, corresponding to *within-patch processing* and *between-patch processing*, respectively. A finer-grained classification, considering emotional aspects, etc., is neither desirable nor necessary. At the same time, note that the proposed processes landscape is a suggestion, which we do claim to be encompassing, but for specific applications, additional processes can be added or omitted.

Limitations The processes landscape was designed with a conventional desktop setup (mouse+keyboard) in mind. Other exploration devices such as mobile phones and tables pose an unequal harder challenge as way less input data can be recorded. Additionally, supporting those devices (if feasible at all) would require extensive extensions to the proposed processes landscape and modelling of the cognitive states. Regardless, the estimation of the cognitive state will always have a high degree of uncertainty, as studies [RF07] have shown that users oftentimes diverge from the interaction patterns commonly observed for a specific state.

5. Conclusion

With the hierarchic processes landscape, we propose a concept of how CISs can be rendered adaptive to address specific user needs and optimize the efficiency of an exploration process. That is, we believe that this collection and classification of interactions/processes/events/states given by literature constitutes a sound bedrock for further developments in the direction of behavior-driven adaptivity. We hope to verify this statement with our future work, which incorporates the implementation of the proposed concept paired with supervised experiments relying on real users.

Acknowledgements

This work was funded by the FWF as part of the project 'Human-Centered Interactive Adaptive Visual Approaches in High-Quality Health Information' (A$^+$CHIS; Grant No. FG 11-B).

References

[ASW06] Arroyo, Ernesto, Selker, Ted, and Wei, Willy. "Usability tool for analysis of web designs using mouse tracks". *CHI'06 extended abstracts on Human factors in computing systems*. 2006.

[BCS16] Battle, Leilani, Chang, Remco, and Stonebraker, Michael. "Dynamic Prefetching of Data Tiles for Interactive Visualization". *Proc. ACM SIGMOD Int. Conf. Manag. Data*. 2016, 1363–1375.

[BOZ*14] Brown, Eli T, Ottley, Alvitta, Zhao, Helen, et al. "Finding waldo: Learning about users from their interactions". *IEEE Transactions on visualization and computer graphics* 20.12 (2014), 1663–1672.

[BvED09] Buscher, Georg, van Elst, Ludger, and Dengel, Andreas. "Segment-Level Display Time as Implicit Feedback: A Comparison to Eye Tracking". *SIGIR*. 2009, 67–74.

[CCAN14] Csikszentmihalyi, Mihaly, Csikszentmihalyi, Mihaly, Abuhamdeh, Sami, and Nakamura, Jeanne. "Flow". *Flow and the foundations of positive psychology* (2014), 227–238.

[CGL*11] Cole, Michael J., Gwizdka, Jacek, Liu, Chang, et al. "Task and user effects on reading patterns in information search". *Interacting with Computers* 4 (2011), 346–362.

[CLWB01] Claypool, Mark, Le, Phong, Wased, Makoto, and Brown, David. "Implicit interest indicators". *International Conference on Intelligent User Interfaces (IUI)*. 2001, 33–40.

[Coo06] Cooke, Lynne. "Is the Mouse a" Poor Man's Eye Tracker"?": *Proceedings of STC*. 2006, 252.

[GA10] Guo, Qi and Agichtein, Eugene. "Ready to buy or just browsing?: detecting web searcher goals from interaction data". *Proceedings of SIGIR*. 2010, 130–137.

[GS00] Goecks, Jeremy and Shavlik, Jude. "Learning users' interests by unobtrusively observing their normal behavior". *Proceedings of the 5th international conference on Intelligent user interfaces*. 2000, 129–132.

[GSL*02] Goldberg, Joseph H, Stimson, Mark J, Lewenstein, Marion, et al. "Eye tracking in web search tasks: design implications". *Proceedings of the 2002 symposium on Eye tracking research & applications*. 2002, 51–58.

[GW09] Gotz, David and Wen, Zhen. "Behavior-driven visualization recommendation". *Proceedings of ACM IUI*. 2009, 315–324.

[HMGO22] Ha, Sunwoo, Monadjemi, Shayan, Garnett, Roman, and Ottley, Alvitta. "A Unified Comparison of User Modeling Techniques for Predicting Data Interaction and Detecting Exploration Bias". *IEEE Trans. Vis. Comput. Graph.* 1 (2022), 483–492.

[HWB12] Huang, Jeff, White, Ryen, and Buscher, Georg. "User see, user point: gaze and cursor alignment in web search". *Proc. ACM Human Factors in Computing Systems (CHI)*. 2012, 1341–1350.

[LC07] Liu, Chen-Chung and Chung, Chen-Wei. "Detecting mouse movement with repeated visit patterns for retrieving noticed knowledge components on web pages". *IEICE Trans. Inf. Syst.* 90.10 (2007).

[MVT16] Mutlu, Belgin, Veas, Eduardo, and Trattner, Christoph. "VizRec: Recommending Personalized Visualizations". *ACM Trans. Interact. Intell. Syst.* 4 (2016).

[OGW19] Ottley, Alvitta, Garnett, Roman, and Wan, Ran. "Follow the clicks: Learning and anticipating mouse interactions during exploratory data analysis". *Computer Graphics Forum*. Vol. 38. 3. Wiley Online Library. 2019, 41–52.

[PC99] Pirolli, Peter and Card, Stuart. "Information foraging." *Psychological review* 106.4 (1999), 643.

[RF07] Rodden, Kerry and Fu, Xin. "Exploring How Mouse Movements Relate to Eye Movements on Web Search Results Pages". *SIGIR Workshop on Web Information Seeking and Interaction* (2007).

[Shu11] Shute, Valerie J. "Stealth assessment in computer-based games to support learning". *Computer games and instruction* 55.2 (2011), 503–524.

[SKS22] Srinivasa, KG, Kurni, Muralidhar, and Saritha, Kuppala. "Stealth Assessment". *Learning, Teaching, and Assessment Methods for Contemporary Learners: Pedagogy for the Digital Generation*. Springer, 2022.

[WVS*21] Weinmann, Markus, Valacich, Joseph, Schneider, Christoph, et al. "The path of the righteous: Using trace data to understand fraud decisions in real time". *MIS Quarterly* (2021).

[YaKSJ07] Yi, Ji Soo, ah Kang, Youn, Stasko, John, and Jacko, Julie A. "Toward a deeper understanding of the role of interaction in information visualization". *IEEE Trans. Vis. Comput. Graph.* 13.6 (2007).