Ecole doctorale SMAER
Sciences **M**écaniques, **A**coustique, **E**lectronique, **R**obotique

Thesis subject 2017

Laboratory : ISIR

University: Université Pierre et Marie-Curie

Title of the thesis: Modeling skill acquisition in Human-Computer Interaction by using Reinforcement Learning

Thesis supervisor: Gilles Bailly

Email contact: gilles.bailly@upmc

Co supervisor: Olivier Sigaud

Collaborations within the thesis:

Program affiliation:

Cotutelle:

University :

This subject can be published on the doctoral school's web site: yes

Thesis's summary (abstract):

This PhD thesis is at the crossroad of Human-Computer Interaction (human behavior) and Artificial Intelligence (Reinforcement Learning). The goal is to model and analyze the behavior of users with graphical user interfaces(?). By modeling human factors, we want to analyze the possible causes of failure explaining why most of the users remain in a local optimum of performance. For instance, explaining why many users continue to navigate in hierarchies of menus instead of using efficient shortcuts (keyboard and gesture shortcuts) even after many years of practice.

We formulate skill acquisition in interactive systems as a Reinforcement Learning problem: users behaviors emerge by finding the optimal policy considering 1) the ecological structure of the interaction, 2) cognitive and perceptual limits of the users and (3) the goal to maximize the trade-off between short-term productivity and long-term performance. Methods will include state-of-art Reinforcement Learning algorithms but should be adapted to take into account the hierarchical and non-stationary nature of the problem. The outcomes of this thesis are threefold: (1) *scientific*. A better understanding of skill acquisition in graphical user interfaces; (2) *Technical*. hierarchical reinforcement learning algorithms dealing with non-stationary armed-bandit problems and memory; (3) *Engineering*. Design tools integrating the models to help designers comparing different interfaces.

Subject

Keywords: Human-Computer Interaction (HCI); Reinforcement Learning; Skill acquisition

Context

The introduction of graphical user interfaces (GUIs) in 1981 played a major role in the democratization of interactive systems in office work, games, medicine, health, finance, etc. Millions of users now spend several hours per day to select commands/applications either on their desktop workstation, smartphone, tablet and now on their wearable devices (e.g. smartwatch). However, GUIs maintain many users in a local optimum of performance. For instance, many users continue to navigate in hierarchies of menus instead of using efficient shortcuts (keyboard and gesture shortcuts) even after many years of practice [Cockburn et al. 2014].
More precisely, GUIs generally rely on two modes. The novice mode relying on menus or toolbar buttons, guides users step-by-step to execute commands. While easy to understand, it only allows users to reach a low level of performance. Conversely, expert mode such as keyboard or gesture shortcuts lets experienced users reach a high level of performance. However, it requires extensive training because it relies on recall rather than recognition. We observe that beginners quickly learn the novice mode, but then fail to achieve a high level of performance because they do not switch to the expert mode. Users continue to use the "novice mode" because they favor short-term productivity rather than long-term efficiency [Caroll and Rosson 1987].
To sum up, **many users are trapped into using slow and repetitive methods even after several years of practice, which has serious and enduring individual (performance/satisfaction) and societal (economy) implications**.

Human-Computer Interaction (HCI) is the science that studies users' behaviors with interactive systems and more precisely, the action-perception loop between one or several users and one or several interactive systems. In HCI, several interaction techniques have been proposed to favor skill acquisition in GUIs relying on advanced feedback [Grosman et al. 2007] and feedforward mechanisms [Malacria et al. 2013]. Several models have also been proposed to explain/predict users' performance such as Fitts' law or the Hick-Hyman Law. However, these models only focus on expert users, i.e. they do not focus on the transition from novice to expert behavior.

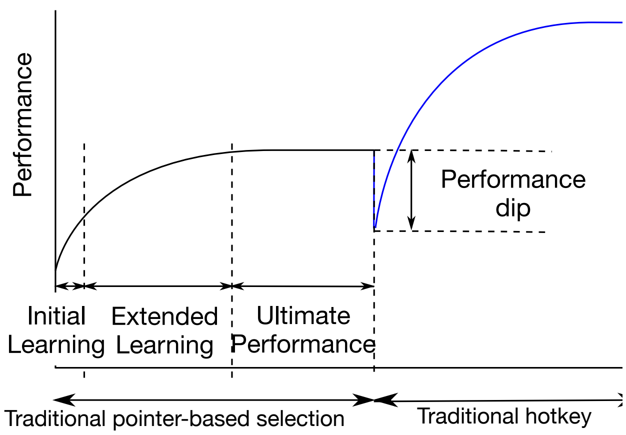

Figure 1: A simple framework of skill acquisition in traditional graphical user interface (GUI) [Cockburn et al. 2014]. Novice users use traditional pointer-based selection methods such as menus or toolbars and quickly reach a low performance ceiling. Many of them are trapped in this local optimum of performance: they do not switch to more efficient methods such as keyboard shortcuts (also called hotkeys) due to a temporary performance dip. Several factors contribute to this performance dip including the memorization efforts required to learn keyboard shortcuts.

Objective

The goal is to model and analyze skill acquisition with graphical user interfaces. By modeling human factors, we want to explain why most of the users remain in a local optimum of performance. While several human factors (e.g. awareness, intrinsic motivations, cognitive abilities, etc.) and system properties (feedback, feedforward, delay, etc.) have been identified [Cockburn et al. 2014], it remains unclear how they impact users' performance due to the complexity of users' behaviors.
Figure 1 shows a simple framework of skill acquisition in graphical user interfaces describing different phases of learning. While the ultimate performance of traditional pointer-based selection methods have been extensively studied, the goal of this thesis is to focus on the modeling of the transition from novice (menu usage) to expert behavior (keyboard shortcuts).

Approach and Methods

The approach builds on the recent advances on Reinforcement Learning [Sigaud&Buffet, 2013]. A key contribution of this literature has been to provide a formal basis for learning an optimal control policy given only a definition of the reward function, the state space, and the action space.

In this thesis, we formulate skill acquisition in interactive systems as a Reinforcement Learning problem. We assume that users behaviors emerge by finding the optimal policy considering:

(1) the ecological structure of the interaction (e.g. how frequently a command is selected),
(2) motor, perceptual and cognitive limits of the users including gaze saccade, gaze fixation duration, mouse movements, intrinsic and external motivations, etc.
(3) the goal to maximize some combination of short-term productivity and long-term performance.

The temporal pattern depicted in Fig. 1 should emerge from solving the reinforcement-learning problem rather than being explicitly given as a users' strategy.

More precisely, we foresee the skill acquisition problem as hierarchical, involving a 2-armed bandit problem at the top level. At this level, the user has to decide which modality (menu or shortcut) to use following an exploration/exploitation strategy. The strategy (or policy) will depend on the expected reward with a modality, i.e. the temporal cost of this modality, the probability to make an error at the time *t* as well as the inclination of the user to adopt an optimal vs. satificing[1] behavior. At the bottom level, the user adopts the optimal strategy (e.g. gaze path, mouse movements) to execute the command with the chosen modality. At this level, the strategy is constrained by the human visual search and motor control system as well as the human short-term/long-term memory system. One challenge to address is how to model users' cognitive processes. Recent machine learning architectures such as [Graves et al. 2016] might be considered.

Originality

The originality of this thesis is twofold. From the HCI perspective, existing models generally focus on experienced users with a single modality rather than focusing on skill acquisition and the transition from novice to expert behaviors in complex interfaces [Bailly et al. 2014]. Moreover, Reinforcement Learning just started to influence the field of Human-Computer Interaction with recent papers such as [Chen et al. 2015] [Chen et al. 2017]. Results show that users' behaviors can be seen as a rational adaptation of the users to the task-related, environmental and psychological constraints, but they but do not address the difficult problem of skill acquisition.

From the Machine Learning perspective, HCI provides an interesting application. Indeed, skill acquisition cannot be considered as a stationary armed-bandit problem because users cannot adopt shortcuts before some practice with menus where shortcuts are displayed. It results that the considered algorithms should incorporate some memory mechanisms, which is not common in reinforcement learning.

Expected results

The outcomes of this thesis are threefold: (1) *scientific*. A better understanding of skill acquisition in graphical user interfaces, i.e. how human factors, system properties and the ecological structure of the interaction impact the transition from novice to expert behavior; (2) *Technical*. Hierarchical reinforcement learning algorithms dealing with non-stationary armed-bandit problems and memory; (3) *Engineering*. Design tools integrating the models to help designers comparing different interfaces.

Work process

During the first year of the thesis, the PhD student will first study existing HCI models (GOMS, ACT-R, EPIC), human factors and system properties (see [Cockburn et al. 2014] for a recent review) involved in the transition from the novice to expert behaviors in interactive systems. The PhD student will also study state-of-art Reinforcement Learning algorithms focusing on those relying on Exploration/Exploitation strategies. S/he will then propose a first model explaining the modalities choice depending on *high-level cognitive considerations* (users' awareness, motivations, attention, etc.). For instance, visual search strategies within the menu will be ignored and menu selection will be considered as a single action independent from the target item.

[1] Satisficing is a decision making strategy that entails searching through the available alternatives until an acceptability threshold is met.

During the second year, the PhD student will refine the models by adding *low-level cognitive considerations*. The PhD student can build on [Chen et al. 2015] regarding visual search strategies (fixation duration, saccade duration, etc.) but should also consider aspects related to motor control (i.e. mouse movements) and how they interact together.

During the third year, the PhD student will integrate the proposed model in a design tool to help HCI designers to create more efficient interfaces. The goal is to inform designers about design choices and to help them quickly compare different alternatives. One challenge will be to optimize the algorithms so that they can provide feedback fast enough.

During the three years, the PhD student will conduct empirical studies in the lab and/or in the field to collect data and evaluate the models.

Environment

The development and the success of our students (from bachelor to Phd) is our highest priority. We provide regular and personal guidance to ensure that students lead successful research projects. The group is part of the exciting and multi-disciplinary ISIR laboratory (robotics, HCI, machine learning, perception, cognitive science, haptics, social interaction, etc.). Gilles Bailly is a young and dynamic researcher with a strong track record at the CHI and UIST conferences (14 papers these last five years and 6 awards), which are the two main avenues of publication in the field of Human-Computer Interaction. Olivier Sigaud has a strong track record in machine learning journals and conferences, and applies his reinforcement learning expertise both in artificial intelligence for robotics and in computational modelling studies.

Profile

Applicants with a strong academic record in any of the two following topic areas are requested to apply: (1) machine learning, (reinforcement learning, Bayesian inference or deep learning) (2) Human-Computer Interaction (skill acquisition, human behaviors, modeling). A Master¹s degree in computer science, cognitive science, or a related area is required.

References

(1) **Gilles Bailly**, Antti Oulasvirta, Duncan P. Brumby, and Andrew Howes. 2014. Model of visual search and selection time in linear menus. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14). ACM, New York, NY, USA, 3865-3874. **Honorable Mention Award**.

(2) John M. Carroll and Mary Beth Rosson. 1987. Paradox of the active user. In Interfacing thought: cognitive aspects of human-computer interaction, John M. Carroll (Ed.). MIT Press, Cambridge, MA, USA 80-111.

(3) Xiuli Chen, **Gilles Bailly**, Duncan P. Brumby, Antti Oulasvirta, and Andrew Howes. 2015. The Emergence of Interactive Behavior: A Model of Rational Menu Search. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15). ACM, New York, NY, USA, 4217-4226.

(4) Xiuli Chen, Sandra Dorothee Starke, Chris Baber, Andrew Howes. 2017. A Cognitive Model of How People Make Decisions Through Interaction with Visual Displays. In Proceedings of the 35rd Annual ACM Conference on Human Factors in Computing Systems (CHI '17). ACM, New York, NY, USA, to appear.

(5) Andy Cockburn, Carl Gutwin, Joey Scarr, and Sylvain Malacria. 2014. Supporting Novice to Expert Transitions in User Interfaces. ACM Comput. Surv. 47, 2, Article 31 (November 2014), 36 pages.

(6) Tovi Grossman, Pierre Dragicevic, and Ravin Balakrishnan. 2007. Strategies for accelerating on-line learning of hotkeys. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '07). ACM, New York, NY, USA, 1591-1600.

(7) Sylvain Malacria, **Gilles Bailly**, Joel Harrison, Andy Cockburn, and Carl Gutwin. 2013. Promoting Hotkey use through rehearsal with ExposeHK. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13). ACM, New York, NY, USA, 573-582.

(8) Steven C. Seow. 2005. Information theoretic models of HCI: a comparison of the Hick-Hyman law and Fitts' law. Hum.-Comput. Interact. 20, 3 (September 2005), 315-352.

(9) Graves, A., Wayne, G., Reynolds, M., Harley, T., Danihelka, I., Grabska-Barwińska, A., ... & Badia, A. P. (2016). Hybrid computing using a neural network with dynamic external memory. *Nature*, *538*(7626), 471-476.

(10) **Sigaud, Olivier** et Olivier Buffet (eds.). *Markov decision processes in artificial intelligence*. John Wiley & Sons, 2013.