Report on the SIGIR 2010 Workshop on the Simulation of Interaction

Leif Azzopardi[1] **Kalervo Järvelin**[2] **Jaap Kamps**[3] **Mark D. Smucker**[4]

[1] University of Glasgow, Scotland
[2] University of Tampere, Finland
[3] University of Amsterdam, The Netherlands
[4] University of Waterloo, Canada

Abstract

All search in the real-world is inherently interactive. Information retrieval (IR) has a firm tradition of using simulation to evaluate IR systems as embodied by the Cranfield paradigm. However, to a large extent, such system evaluations ignore user interaction. Simulations provide a way to go beyond this limitation. With an increasing number of researchers using simulation to evaluate interactive IR systems, it is now timely to discuss, develop and advance this powerful methodology within the field of IR. During the SimInt 2010 workshop around 40 participants discussed and presented their views on the simulation of interaction. The main conclusion and general consensus was that simulation offers great potential for the field of IR; and that simulations of user interaction can make explicit the user and the user interface while maintaining the advantages of the Cranfield paradigm.

1 Introduction

The use of simulation to evaluate retrieval systems has a long history in IR, especially before the availability of large-scale test collections developed in the 1990s [6, 20]. In recent years, simulation has mainly been used to overcome limitations of traditional test collections, in particular to evaluate adaptive or interactive IR [e.g., 22]. The different types of experiments that can be performed to examine Interactive IR may be classified into four classes [14, p.210]:

1. observing users in real situations (real users; no simulation);
2. observing users performing simulated tasks;
3. performing simulations in the lab without users (simulation of interaction; no users); and
4. traditional lab research (no users and no simulation).

Interactive IR may be studied experimentally with real searchers performing real or simulated work tasks (class 1 and 2 respectively). However, these types of experiments require a lot effort to setup in order to conduct reliable studies, and they tend to be tedious and