**Title**
Computing the Canonical Subset of User Protocols

**Permalink**
https://escholarship.org/uc/item/0ts863g7

**Journal**
Proceedings of the Annual Meeting of the Cognitive Science Society, 31(31)

**ISSN**
1069-7977

**Authors**
Bogunovich, Peter
Mankowski, Walter
Salvucci, Dario
et al.

**Publication Date**
2009

Peer reviewed

# Computing the Canonical Subset of User Protocols

**Walter Mankowski**
Drexel University

**Peter Bogunovich**
Drexel University

**Ali Shokoufandeh**
Drexel University

**Dario Salvucci**
Drexel University

**Abstract:** A common problem in cognitive science research is the large volume of behavioral protocol data recorded during the execution of the tasks being studied. The analysis of these large data sets has often been a tedious and time-consuming process, and automated analysis methods have been slow to develop. We have developed an automated method to find canonical behaviors: a small subset of protocols that is most representative of the full data set, providing a "big picture" view of the data with as few protocols as possible. The method takes advantage of recent algorithmic developments in computational vision, which we have adapted to the comparison of behavioral protocols. No a priori model is required, just a similarity measure between pairs of behaviors. Initial experiments show that canonical sets of web-browsing protocols found by our method compare well with those found by expert human coders.