

Analyzing Process Data from Game/Scenario-Based Tasks: An Edit Distance Approach

Jiangang Hao
Educational Testing Service
Princeton, New Jersey 08541
jhao@ets.org

Zhan Shu
Educational Testing Service
Princeton, New Jersey 08541
zshu@ets.org

Alina von Davier
Educational Testing Service
Princeton, New Jersey 08541
avondavier@ets.org

ABSTRACT

Students' activities in game/scenario-based tasks (G/SBTs, hereafter) can be characterized by a sequence of time-stamped actions of different types with different attributes. For a subset of the G/SBTs where only the order of the actions are of great interest, the process data can be well characterized as a string of characters (action string, hereafter) if we encode each action name as a single character. In this paper, we report our work on evaluating students' performances by comparing how far their action strings are from the action string that corresponds to the best performance, where the proximity is quantified by the edit distance between the strings. Specifically, we choose the Levenshtein distance, which is defined as the minimum number of insertion, deletion and replacement needed to convert one character string to another. Our results show a strong correlation between the edit distances and the scores obtained from the scoring rubrics of the WELL task from NAEP TEL, implying the edit distance to the best performance sequence can be considered as a new feature variable that encodes information about students' proficiency, which sheds light on the value of data driven scoring rules for test and task development and for refining the scoring rubrics.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Copyright 20XX ACM X-XXXXX-XX-X/XX/XX ...\$15.00.