1. INTRODUCTION

Fundamentally, a graph is a simple concept. At a basic level a graph is a set of relationships \( \{e(n_0,n_2), e(n_0,n_3), \ldots, e(n_{j-1},n_j)\} \) between elements. This simple concept, however, has afforded the development of a complex theory of graphs [1] and rich algorithms for combinatorics [7] and clustering [4]. This has, in turn, made graphs a fundamental part of educational data mining.

Many types of data can be naturally represented as graphs such as social network data, user-system interaction logs, argument diagrams, logical proofs, and forum discussions. Such data has grown exponentially in volume as courses have moved online and educational technology has been incorporated into the traditional classroom. Analyzing it can help to answer a range of important questions such as:

- What path(s) do high-performing students take through online educational materials?
- What social networks can foster or inhibit learning?
- Do users of online learning tools behave as the system designers expect?
- What diagnostic substructures are commonly found in student-produced diagrams?
- Can we use prior student data to identify students’ solution plan, if any?
- Can we use prior student data to provide meaningful hints in complex domains?
- Can we identify students who are particularly helpful based upon their social interactions?

Thus, graphs are simple in concept, general in structure, and have wide applications for Educational Data Mining (EDM). Despite the importance of graphs to data mining and data analysis there exists no strong community of researchers focused on Graph-Based Educational Data Mining. Such a community is important to foster useful interactions, share tools and techniques, and to explore common problems.

2. GEDM 2014

This is the second workshop on Graph-Based Educational Data Mining. The first was held in conjunction with EDM 2014 in London [17]. The focus of that workshop was on seeding an initial community of researchers, and on identifying shared problems, and avenues for research. The papers presented covered a range of topics including unique visualizations [13], social capital in educational networks [8], graph mining [19, 11], and tutor construction [9].

The group discussion sections at that workshop focused on the distinct uses of graph data. Some of the work presented focused on student-produced graphs as solution representations (e.g. [14, 3]) while others focused more on the use of graphs for large-scale analysis to support instructors or administrators (e.g. [18, 13]). These differing uses motivate different analytical techniques and, as participants noted, change our underlying assumptions about the graph structures in important ways.

3. GEDM 2015

Our goal in this second workshop was to build upon this nascent community structure and to explore the following questions:

1. What common goals exist for graph analysis in EDM?
2. What shared resources such as tools and repositories are required to support the community?
3. How do the structures of the graphs and the analytical methods change with the applications?

The papers that we include here fall into four broad categories: interaction, induction, assessment, and MOOCs.
Weber & Vas [21], takes a different tack and focuses not on graphs when learning programming. He argues that the conceptual work Hicks et al. focuses on the use of path weighting to select impact of social relationships on users' class performance. identifying stable sub-communities of users and of assessing the not on local behaviors but on social networks with the goal of Clayphan present work on the use of graph-based metrics to focus on the unique problems posed by MOOCs. Vigentini and based self-assessment that is designed to bring these conceptual relationship graph affords a better mechanism for automatic as- questions and in giving appropriate assignments. Guerra takes support content developers and educators in examining a set of representing solutions or interactions but on relationships. Leo- John et al. present a novel approach for identifying closely-related word problems via semantic networks. This work is designed to support content developers and educators in examining a set of questions and in giving appropriate assignments. Guerra takes a similar approach to the assessment of users’ conceptual changes when learning programming. He argues that the conceptual relationship graph affords a better mechanism for automatic as- sessment than individual component models. This approach is also taken up by Weber and Vas who present a toolkit for graph-based self-assessment that is designed to bring these conceptual structures under students’ direct control.

And finally, Vigentini & Clayphan [20], and Brown et al. [2] focus on the unique problems posed by MOOCs. Vigentini and Clayphan present work on the use of graph-based metrics to assess students’ on-line behaviors. Brown et al., by contrast, focus not on local behaviors but on social networks with the goal of identifying stable sub-communities of users and of assessing the impact of social relationships on users’ class performance.

4. REFERENCES


