Abstract

While technological progress made more content readily available, the information load makes it hard to navigate the content space and attain learning goals. The exponential growth of specialized knowledge challenges learners, scholars, teachers, and educational institutions. For novice learners, knowing where to start and what courses contribute towards educational goals, and in what sequence, are key problems. For scholars and researchers, keeping abreast of latest developments in their field, embarking on new or interdisciplinary research, or gaining deeper understanding of the foundations of a field, all require understanding of how the knowledge space is structured and connected. For educators and educational institutions, developing curricula of proper coverage and design implies knowledge of content mapping, overlap, and dependency. For emerging online learning organizations, course recommendation and personalization of the learning experience require complex learner-content modeling. As the space of knowledge grows in scope and connections, new questions arise: How do we define coherent modules’ or topics in this ever growing space? How do we represent content at different levels of abstraction? Can we discover latent patterns and relations in the content space? What are the characteristics of well-designed curricula? Can we formalize and guide the personalization of learning paths? Can we automate curriculum planning and generation? What is common between all these questions is the their relevance to content. In this thesis, we propose to model content at the concept, module, and curriculum levels which we define in detail in this proposal. At each of these levels, we focus on three main problems: representation, (planning and) generation, and evaluation. We propose models, methods, and evaluation criteria and metrics for the specific tasks we define to tackle these main problems. The representation, data structures, results and resources at each level serve to solve problems in the next level. Ultimately, the goal is to fully automate the navigation of the knowledge space given specific educational goals (coverage) and constraints defined on the learner's resources.