

Research Statement

Amanda M. Holland-Minkley
hollandm@cs.cornell.edu

In my graduate work, I have investigated the presentation of formal mathematics through natural language, primarily focusing on the organizational and planning issues involved in structuring a mathematical text. I have implemented a prototype system for converting computer-generated proofs into textbook-style English texts via an empirically-motivated intermediate plan representation. My focus has been on the discourse planning component, with the goal of connecting highly structured formal proofs with the text structure required for effective communication via a mathematically- and language-coherent intermediate representation. Doing so required drawing on work from natural language generation, knowledge representation, automated reasoning and formal mathematics. To ground the work, I performed a series of studies of how people knowledgeable in formal mathematics prefer to represent and manipulate this mathematics in a textual form. The resulting system is able to produce proof texts that are competitive with human-produced proofs in their effectiveness at communicating proof content.

My published work and system implementations to date have all operated under a single user model representing an average current user of our automated reasoning system. However, it is clear both from our study of math texts and from related work on knowledge representations and language generation that a truly robust system must be able to tailor itself to an individual users' knowledge and textual preferences. It is natural to represent the facts about which a system can communicate as a hierarchy or tree in which we can mark the boundaries of a users' knowledge rather than exhaustively testing and listing the facts familiar to each user. Many pedagogical tools assume there is such a natural ordering, and treat the teaching task as helping students master the entirety of the course content hierarchy.

In the specific case of our reasoning assistant, we have an extensive directed dependency graph between all established mathematical facts within our system. This complex graph alone cannot be used as-is as a hierarchy of knowledge. However, I believe that a suitable hierarchy could be inferred from this graph. I suggest that, as we were able to examine the structure inherent in a formal proof and identify mathematical communication conventions to transform that structure into a textual proof plan, a similar strategy grounded in empirical observations of users' relevance judgments will be effective in transforming our dependency graph of math objects into a hierarchy of knowledge. If this can be done effectively, the same technique for inferring a hierarchy of knowledge should be applicable to the wide range of structured databases of knowledge existing in the many expert systems in development over a wide range of topics.

Though my work to date has only been tangentially related to the human-interfaces community, I have become interested in connecting my work more directly to that field. Through the variety of human-subject studies I have performed to determine the style of texts with which users prefer to interact, I have observed many openings for customizing automatically generated texts not just through their surface content but also through the tools available with which to interact with the information represented in the text. In the setting I have been exploring, where the natural language texts produced are necessarily summarizing a much richer underlying content, interfaces allowing users to see the formal content justifying an English proof-sentence, search for similar statements in other proof-texts, or regenerate proof texts with a request for elaboration in particular areas are desired. In this setting, the generated texts become simply another customizable component of the users computing environment.

There are many pieces of these projects that could be explored usefully by students of even an intermediate undergraduate experience. These projects would be suitable for undergraduates interested in research or independent study projects. I would also be enthusiastic about working with faculty and/or students from related fields, such as psychology, who may wish to explore those cognitive sciences connections.