Hypatheon: A Mathematical Database for PVS Users

Extended Abstract

Ben L. Di Vito (benedetto.l.divito@nasa.gov)
NASA Langley Research Center

Researchers and practitioners of theorem proving-based formal methods have a great need for formalized mathematical knowledge. This community pursues the goal of assuring proper operation of critical computing systems. Successful engineering applications are increasing despite their high cost.

Consider, for example, the formal verification of key properties of air traffic management algorithms, a current line of research at NASA Langley. This work emphasizes rigorous proofs using an interactive theorem proving tool (PVS). Models are developed that capture geometric scenarios constrained by the physics of motion. Proofs in this area hinge on trigonometry, calculus and other theories of continuous mathematics. Although these are standard parts of engineering mathematics, in common use for over a century, much of it has yet to be codified in a form that supports mechanical theorem proving.

Tools such as PVS have some excellent capabilities, but due to technological immaturity, are woefully "under-educated," typically having the mathematical knowledge of a middle school student. They force us to spend too much time teaching them and make us repeat the exercise too often. These circumstances create a large drag on productivity and will limit the uptake of formal methods by engineers. One way to characterize the problem is that we need to "send PVS to high school," plus one or two years of college, before it will be adequately knowledgeable to serve the needs of formal methods practitioners.

Thus, we have two problems: collecting formal mathematics and disseminating relevant theories to users. Experience with deductive techniques has shown that fully formal axiomatizations generate large numbers of definitions and supporting lemmas. Given the substantial breadth of engineering mathematics, we estimate that to codify a sufficiently rich portion of it would require millions of definitions, theorems and other deductive artifacts. Since few of us have the eidetic powers needed to recall millions of named items and reliably choose among them, only careful organization and automated search would enable us to exploit such a body of knowledge.

At NASA Langley, we are developing a mathematical database for PVS together with a set of supporting services, which has been named Hypatheon (for Hypatia and her father Theon). This effort is an outgrowth of our long-standing work in developing PVS libraries. By investing in infrastructure for deductive knowledge, we hope to attract contributions from the PVS community. If users benefit from what we offer, we expect many will be motivated to reciprocate, and a passive collaboration process should emerge, which will lead, over time, to a comprehensive collection of artifacts usable by designers of dependable computing systems.

We wish to publish deductive/mathematical knowledge by hosting a dedicated

Web server and providing a specialized set of services to PVS users. These services include: 1) a Web-based interface mechanism to issue queries against the database and accept submissions of new content for inclusion in the database; 2) a client module to complement PVS, offering proof-side assistance during prover sessions and automating the discovery and acquisition of relevant theorems; and 3) an extensible platform for implementing future services based on a programmatic interface.

A prototype server is operating internally at NASA Langley. The basic capabilities now operational include a Web service that supports browser access to the database, a workflow subsystem that enables automatic submission of new content (in the form of theories in the PVS language), a PVS client that supports proof-side queries, and a mechanism for this client to download new libraries/theories on demand.

Initial database content has been created from existing PVS libraries, many of which were developed at NASA Langley. This has resulted in a core dataset having roughly 20 libraries, 450 theories, 1200 functions, and 4000 theorems. Later we expect to add libraries for domain-specific formalizations such as fault tolerant computing and air traffic management algorithms. Populating the database further will depend on success in encouraging PVS users to submit new content.

Several pragmatic aspects of the overall concept and design are still under study, such as status or maturity indicators for libraries, coexistence of multiple library versions, propagation of revisions through the database, submission policies, maintainer responsibilities, and the addition of editor role(s), i.e., people who are responsible for content in various areas, who decide what to accept, who work with submitters/maintainers, etc.

Our next step in query development aims for greater automation. After devising heuristics for ranking search results, we will try to select suitable lemmas automatically or at least narrow the candidates to a manageable number. Later we will add proof objects to the database. This will allow us to relate proof steps to other lemmas or proofs, which in turn will help identify patterns and idioms, support searches based on proof content, and enable proof cloning.

Our near-term plans are to continue to refine the current prototype and prepare for a public server rollout in early 2004. Performance goals for the server are to support 1000 libraries, 10,000 theories, 100,000 function definitions, and 1 million theorems (formulas). We expect to develop advanced query capabilities, add proof handling features, and pursue data mining opportunities.

Although Hypatheon is intended primarily for engineering use, a relationship with the mathematics community clearly would be valuable. Mathematicians or their students might wish to submit content or guide the organization of Hypatheon's knowledge. Naturally, such participation would be quite welcome. Some mathematicians might eventually find it desirable to construct fully formal proofs, in which case our services could be of interest, along with others under development by MKM researchers. Our presentation for NA-MKM 2004 offers a brief overview of the Hypatheon project, the tools under development, and our plans for continuing research and development.