FROM HOURS TO SECONDS: A CRISIS AVERTED

Summary

The customer is a specialist in the field of data governance with a particular emphasis on the management of medical information via a SaaS product.

This customer is using Haskell to load and pre-process a large body of fixed information that defines the terminology framework for medical records. Haskell is a good choice for this due to its high performance and strong compile-time guarantees on the correctness of algorithms while still supporting exploratory programming in an interactive environment.

Before the Haskell code was finished the customer was facing an important deadline with a very tight timescale. The customer's engineers were having problems completing this component. In particular, the program was unable to finish processing when given several million records of input..

Corporation Type: SaaS Software Vendor
Industry/Sector: Medical knowledge management
Project Type: Import and processing of medical terminology frameworks
Technology Used: Haskell, Mac

Project Requirements

The primary requirement was to fix the import process so that the terminology framework could be converted to a different format and used in the SaaS product in time for the deadline.

Secondary requirements were to review the existing code to suggest improvements and to mentor the customer's engineers in more advanced Haskell techniques.

The Solution

FP Complete's experience with Haskell and our ability to analyze complex problems enabled us to fix this well before the deadline. We discovered that the import process was extremely slow due to some exponential algorithmic complexity deep inside the program. We changed it to use an index instead of a linear lookup and the runtime went down from 25 hours to under 3 seconds.

In the remaining time, we assisted with refactoring and general cleanup of the code, and coaching the engineers via series of video calls with screen sharing. We also wrote a small proof-of-concept project that outlines how the software might be rearchitected to provide a much more intuitive and performant foundation for ongoing development.

New Challenges for FP Complete

This project involved several challenges:

- Gaining rapid understanding of a new specialist field (medical ontologies).
- Writing a tool to automate long-running benchmarks of code that may not terminate.
- Writing a tool to analyze performance degradation as a function of input size.
- Enhancing training materials on pitfalls to avoid and best practices to follow.

Conclusion

FP Complete was able to make an immediate and significant difference to a key component of the customer's software stack, despite having very little time to come up to speed on a complex codebase in a specialized business domain. The data was imported and processed before the deadline, and the codebase is ready to handle imports with ease in the future.

We were also able to equip the customer with the ability to maintain and extend their code going forward, and provide them with techniques for writing code that is both performant and correct.