

Out of the Tar Pit

Ben Moseley
ben@moseley.name

Peter Marks
public@indigomail.net

February 6, 2006

Abstract

Complexity is the single major difficulty in the successful development of large-scale software systems. Following Brooks we distinguish *accidental* from *essential* difficulty, but disagree with his premise that most complexity remaining in contemporary systems is essential. We identify common causes of complexity and discuss general approaches which can be taken to eliminate them where they are accidental in nature. To make things more concrete we then give an outline for a potential complexity-minimizing approach based on *functional programming* and *Codd's relational model of data*.

1 Introduction

The “software crisis” was first identified in 1968 [NR69, p70] and in the intervening decades has deepened rather than abated. The biggest problem in the development and maintenance of large-scale software systems is complexity — large systems are hard to understand. We believe that the major contributor to this complexity in many systems is the handling of *state* and the burden that this adds when trying to analyse and reason about the system. Other closely related contributors are *code volume*, and explicit concern with the *flow of control* through the system.

→ Kontext řešeného problému, motivace

The classical ways to approach the difficulty of state include object-oriented programming which tightly couples state together with related behaviour, and functional programming which — in its pure form — eschews state and side-effects all together. These approaches each suffer from various (and differing) problems when applied to traditional large-scale systems.

→ Stručné shrnutí dřívějších přístupů

We argue that it is possible to take useful ideas from both and that — when combined with some ideas from the relational database world — this approach offers significant potential for simplifying the construction of large-scale software systems.

→ Podstata práce, k čemu směřuje

The paper is divided into two halves. In the first half we focus on complexity. In section 2 we look at complexity in general and justify our assertion that it is at the root of the crisis, then we look at how we currently attempt to understand systems in section 3. In section 4 we look at the causes of complexity (i.e. things which make understanding difficult) before discussing the classical approaches to managing these complexity causes in section 5. In section 6 we define what we mean by “accidental” and “essential” and then in section 7 we give recommendations for alternative ways of addressing the causes of complexity — with an emphasis on avoidance of the problems rather than coping with them.

In the second half of the paper we consider in more detail a possible approach that follows our recommended strategy. We start with a review of the relational model in section 8 and give an overview of the potential approach in section 9. In section 10 we give a brief example of how the approach might be used.

Finally we contrast our approach with others in section 11 and then give conclusions in section 12.

→ Struktura práce vzhledem k cílům