

Changing Technical Writing from a Craft to a Science

Introduction

Some documents have huge leverage, whether they are specifications for multi-billion dollar projects, or legislation governing decades of expenditure. This leverage suggests we should take more care in their creation. Currently, the creation of technical and legal documentation – specifications, written plans, technical and scientific papers, laws – is still a craft. We have tools to help us with the formatting and the spelling, but we have no tools to help us with the meaning. As long as the salient points are touched on and there is nothing that is garbled at close range, the writing is acceptable. There is no attempt to quantify whether the text is ambiguous, inconsistent or confusing, or will exceed the cognitive limitations of the reader. High value documents are typically built incrementally by many hands, not tossed off by a lone hand in a single sitting. This incrementalism allows many errors to occur. If we can create an automated reader, we can use it to find errors and inconsistencies that would otherwise slip through.

Why Hasn't This Been Done Already

There are some problems that are well suited to a reductionist approach, and there are other problems – systems problems - where a reductionist approach moves the solution further away. Splitting the processing of text into sub-problems involving grammar and semantics, while seemingly simplifying it, pushes the problem further away, as should be obvious from observing the operation of such a system on the first half dozen sentences. Processing of text requires a cognitive approach, where reductionism doesn't work, and phrasing cannot be predefined. This means everything we know about programming has to be thrown away.

Let's use an aircraft as an example. When the Wrights came to build an aircraft, they found the science of aerofoils to be wrong, so they had to do it again for themselves, they found that the available engines were built for terrestrial use, so they had to make a new, much lighter one. If a system operating in a new way is desired, it is likely that the science is all wrong, either because of concentration on a facet, or it doesn't scale up, or because it has not been tested in the environment in which it will be used.

Technical text represents a synthesis of many things –

- Propositional Logic
- Temporal Logic
- Existential Logic
- Locational Logic
- Relations on Relations
- Groups and Sets
- Persistent States
- Similarity
- Document Structure

Nobody has seriously attempted a synthesis of the different logics, or the attempts have been so clumsy as to be worthless. One reason is that natural language is

already a very good synthesis over all of logic, so an opaque and limited notation would have little to recommend it to anyone other than its originator.

Many people have worked on the logics separately, but have assumed everything else is constant, resulting in a simpler structure for their purpose, but also ensuring their creation cannot be part of a synthesis.

We have a synthesis across all these things, so can analyse documents in a way not previously possible - the errors and inconsistencies in documents can be found and pointed out, so others can rectify them. Many errors are laughable when pointed out, yet would live in the documents and cause confusion for decades without automatic analysis.

This isn't meant to be a criticism of technical writers - all humans have a limit of about four pieces of new information in play. Give them a useful tool and the quality goes up - it is very hard to draw a perfect circle freehand, but with the right tool it is easy.