

The Business-Aware Road to Strong AI

Petr Baudiš (pasky@ucw.cz)

Spring 2014, r1

There have been many attempts to build a strong AI in the past, and many more are ongoing. Often, they get captured in a variety of fallacies — strictly formal logic approaches, insisting on a robotic embodiment, or focusing on extreme human imitation like emotions or copying neural circuits and the brain architecture.

The goal of creating an intelligence on or surpassing human level, and working and behaving similarly to human intelligence, is still far away and we have little idea how to get there. Instead of setting this goal straight away, I propose an *incremental* route to the strong AI built around practical applications. The route to strong AI is long and therefore the development should have a truly sustainable funding model, built around a self-reinforcing profit feedback loop. This means that each milestone along the way must be in itself an obvious business opportunity rather than just a cute prototype.

Where to start with this requirement in mind? In practice, a synergy of multiple fields will emerge in time, but I think that the problem of *knowledge* is more important and easier to solve than the problem of *behavior*. Simply put, I claim that average(!) humans are much better at learning and improvising a behavior than grokking vast knowledge bases and using them productively. At the same time, we have vast amounts of knowledge in digital form that we can feed our programs, but digital data on behaviors are currently much harder to get.

Therefore, rather than, say, building smarter and more tractable robots, I claim that the lowest hanging fruit in computational intelligence is knowledge processing. More specifically, natural language processing as most of the interesting knowledge is in the form of unstructured English text. I'm not alone with this vision; it seems that e.g. Google and IBM are looking in exactly the same direction!

The key to “smart computer behavior” is statistics and for statistics to work reliably, big data is necessary. We have a vast corpus of most human knowledge at our fingertips with the English Wikipedia, and the costs are minimal. I believe that with the right algorithms and enough computational power (affordable by now), we can convert the Wikipedia to a sort of post-processed “memory” that can be used to reason about the concepts of worlds both real and virtual. Most even trivial concepts are actually spelled out and the rest is still embedded in the texts in a latent form.

The same algorithms can then be used for processing other texts; law documents and codexes, engineering datasheets, computer manuals, medical literature, perhaps even books of fiction; concepts discovered in these documents can be related and contextualized using the learned “Wikipedia memory”.

The most direct application of such features is the ability to *answer questions* about the world or a specific resource. Obviously, this field is AI-complete as the questions can grow very

Road to AI

complex;¹ but even less powerful implementations would be satisfying. The first stage is getting the computer to answer trivia questions reliably enough, much like a simpler version of IBM Watson.

In terms of business, I see applications in science, engineering, law, user support and interaction, even journalism. I'm still thinking the exact details through, and if you have a cool application idea, please do share it.

In terms of AI development, this can serve as a “memory module” for a reasoning, goal-oriented action-taking system as well as a platform for various specialized task solvers. If we are building a system that would solve a real-world physics problem (e.g. from entrance exams), we can use this module to convert real-world concepts to a pure symbolic problem description, clarifying both the variables and the processes in play.

Initially, the AI I envision will not be embodied, it will not have emotions or desires, it won't work “like an artificial brain” or “artificial being”. Its living environment will be filesystems, network ports and internet chatrooms. Its memory will be swathes of post-processed natural text structures, seeded by Wikipedia and few ontological databases. It won't be able to iron your clothes or write poems, but it will be able to help you with engineering problems or do case law reviews. Then, we can improve.

Join us at the Brmson project.

¹All the way to “What is the meaning of life, universe and everything?”