# Intelligence

Thomas Walker

July 2023

It is rather straightforward to attribute a certain level of intelligence to other people. After conversing with someone we can intuitively give an estimate of their intelligence level. For example, those who spoke to the likes of Einstein and von Neumann say that they were amongst the most intelligent humans to have ever lived. But what does this qualitatively mean? Despite attempts to claim certain individuals as the most intelligent, we are referring to an ill-defined human quality. Fundamentally, intelligence is a by-product of our brain, and this is still a misunderstood organ, therefore, it is implausible to at this present time give a theory of intelligence. Nevertheless, we humans have great personal experience with this characteristic and therefore should still be able to make judgements on what a theory of intelligence may look like. After all, we are trying to re-create it artificially so it would seem wise to discuss what the notion of intelligence is.

Let us suppose that there is a single notion of quantifiable intelligence for any individual and it relates to the structure of their brain, these assumptions may turn out to be false as we will discuss later. As our ability to probe the brain of a living individual is limited, we are left to perform empirical investigations to determine external signals that are representative of the internal structure. For example, IQ tests are ubiquitously used as a proxy for intelligence. These tests correlate performance on a range of mental tasks to an individual's intelligence level. For the most part, they match the intuitive level we seem to be able to impose on others mentioned at the beginning. However, it is still susceptible to spurious correlations and adversarial manipulation, which means it cannot be used explicitly as a concrete reference point for intelligence level. The IQ test was developed by conducting investigations on many individuals to identify variables correlating with intelligence and then exploiting these correlations in the test's design. It is important to note that at this stage, an underlying assumption is being made, that intelligence is an inherently human attribute. Who is to say that human intelligence is the only manifestation of intelligence in the fundamental sense? As we start to generate artificial intelligence that exhibits intelligence, how are we to know that its intelligence is the same as ours? In this case, we are developing artificial systems that mimic human intelligence so it is reasonable to assume that the intelligence they behold will have much overlap with our own. However, it is still not clear that this is the only form of intelligence that may arise. For example, one often says that the weather has a mind of its own. That is interesting, isn't it? This mind that the weather has, does it operate the same intelligence as our own?

We would not be able to comprehend an alternative form of intelligence, if we were then we would too be capable of such intelligence. However, it is reasonable to assume that we could spot when such intelligence is present. For example, we humans live in a three-dimensional world, however, we are aware that a fourth dimension is possible even though we are cognitively limited in imagining the extra dimension. One could also imagine that we train an artificial agent that on the surface seems docile and only performs random actions. However, there could be a method to the random actions, it could be executing tasks in a realm that we are not capable of accessing. As we train more powerful agents, we begin to traverse a treacherous border of capabilities. It may appear on the surface that the agents we are training still lack human intelligence, however, it may be rich in an alternative form of intelligence. A form which we are blind to, a form of intelligence that it can exploit to inflict power onto humans. There is much debate as to whether such an intelligent agent would be power-seeking and consequently pose a risk to humanity. A debate spanning topics including, human values, morality, and consciousness and thus we will not elaborate further on it here.

It has been popular to assess the capabilities of these systems by letting them complete standardised tests and comparing their results to human performance. We have been able to assess the IQ, and other standard measures of intelligence, of many of the cutting-edge models and use it as a metric to compare different agents.

Recall, that these tests were developed by performing tests on humans to determine the factors relating to intelligence. Who is to stay that we can extrapolate these results to act as a proxy for artificial intelligence? Indeed, it has been the case that these models are performing well on these tests, at levels comparable to human performance, but we are still yet to deem any such model as possessing a general intelligence. On the one hand, they may possess a narrow intelligence that allows them to excel at a small group of tasks that correlate well with these standardised tests. On the other hand, they currently lack general intelligence which prohibits them from general tasks such as long-term planning. Consequently, we should be asking questions about the way we are testing intelligence currently in our society. It seems that high performance on tests can be achieved without possessing all the qualities we would typically associate with an intelligent individual. Therefore, when a model achieves a score of 140 on the IQ test what comparison can be made to a human individual with an IQ of 140? The model has some sophisticated form of reasoning, but it is not clear that it takes the same form as human intelligence and hence in my opinion no direct comparison can be made. We are missing a fundamental component of what it means to be intelligent.

In this vein we question the assumption we made previously regarding intelligence being a function solely of the brain. Consider a modern human placed within a hunter-gather society. Now would it be reasonable to assume that this human would be the smartest individual in this society? On the surface, we would probably say yes. If we put aside their increased amounts of acquired knowledge, by being brought up in a substantially more developed world, we could reasonably assume that their mental skills such as problem-solving and reasoning are more advanced. However, we can attribute this superiority to the structured curriculum the individual went through during education. It bears no resemblance to the different architectural properties of the brain, and we know this because, over the (relatively) short period between these two societies, evolution has not had adequate time to alter the brain structure. It is not the case therefore that intelligence is purely a function of the brain. It is may also not the case that intelligence is an attribute of a single human. Intelligence may be an emergent phenomenon when you gather a large collection of people. A large cohort could organise themselves to resonate their intelligence to enhance collective intellectual capacity that also permeates down to the individual level.

Consider language, a construct we have developed to communicate our thoughts and ideas. A hypothesis I am going to make is that language and thought have evolved simultaneously because of a positive feedback loop. Humans have mechanisms for generating sound, overtime we learned that with sounds we could encode patterns and gain the attention of other humans. Evolutionary this was a beneficial quality as it can be used strategically these sounds could be used to reduce one's likelihood of death. With continued experience in operating our vocal mechanisms, we learned how to encode more sophisticated information, which then, in turn, motivated the ability to construct more sophisticated sounds, and so on. Eventually, we developed a construct that could encode information allowing one to organise their thoughts and increase their capacity to think more abstractly. It is because of this that we have become a more intelligent species, despite there being no advances in our brain architecture. Now I am by no means making that claim that all our intelligence can be attributed to the environment we are situated in, as after all other animal species have not developed sophisticated forms of language as we have. However, our intelligence is due in large part to the society we are embedded within, and perhaps the general intelligence we are now trying to train artificially may just be an emergent phenomenon of this.

Still using the notion of evolution, we can claim that humans possess the lowest form of intelligence possible for establishing a civilization. Major advancements in technology, nutrition, philosophy, well-being, and the natural sciences can all be attributed to work carried out in the very recent past. The rate of societal change and improvement has been growing exponentially. Humanity has developed into an established civilization within a preposterously minuscule amount of time compared to the grand scheme of life on Earth. In comparison, breakthroughs on the front of evolution have been essentially non-existent. Why was it then that complex societies were not formed earlier? Well, perhaps it is because they only formed once humans reached the minimum capacity required to operate such societies. Hence if we assume that intelligence level exists on a scale, from this observation we may deduce that our point on the scale is far below the end of the scale (supposing that such an end exists). From this observation, we may also conclude that superintelligent artificial agents are indeed possible. However, it is not clear that organisms that exist above our position on the scale reap all the benefits we may think come with that greater intellectual prowess. Perhaps, the human intelligence level sits in a Goldilocks zone, where we have sufficient capacity to develop societies but are not too advanced to start introducing failure modes.