In the book *Superintelligence*, Nick Bostrom defines a superintelligence as "any intellect[1] that greatly exceeds the cognitive performance of humans in virtually all domains of interest."

The "virtually all domains" part of this definition is important: while current AIs are often much more capable than humans within some narrow domain — e.g., chess — *superintelligence* is usually used to mean a system that far exceeds human intelligence in most domains.

Alternate phrasings

- What is ASI?

Related

- What are the differences between AGI, transformative AI, and superintelligence?
- What is "transformative AI"?
- What is artificial general intelligence (AGI) and what will it look like?
- Why might we expect a superintelligence to be hostile by default?

Scratchpad

- definition, existence of Bostrom's book
- maybe mention that superintelligence in theory doesn't have to be AI but that's the version that's usually talked about?
- some specific important domains?
- probably not a far future thing (at least it being very powerful does not imply it being far in the future)
- something about there being a lot of room above humans
- how this relates to risk (we'd lose to a superintelligence by default, no do-overs); maybe also benefits

The LW tag

A **Superintelligence** is a being with superhuman intelligence. Specifically, Nick Bostrom (1997) defined it as

[1] "Superintelligence" usually refers to an *artificial* intelligence, but it doesn't have to. For example, a human whose intelligence was technologically augmented to a sufficient degree could be superintelligent.

> "An intellect that is much smarter than the best human brains in practically every field, including scientific creativity, general wisdom and social skills."

The Machine Intelligence Research Institute is dedicated to ensuring humanity's safety and prosperity by preparing for the development of an Artificial General Intelligence with superintelligence. Given its intelligence, it is likely to be incapable of being controlled by humanity. It is important to prepare early for the development of friendly artificial intelligence, as there may be an AI arms race. A strong superintelligence is a term describing a superintelligence which is not designed with the same architecture as the human brain.

An Artificial General Intelligence will have a number of advantages aiding it in becoming a superintelligence. It can improve the hardware it runs on and obtain better hardware. It will be capable of directly editing its own code. Depending on how easy its code is to modify, it might carry out software improvements that spark further improvements. Where a task can be accomplished in a repetitive way, a module preforming the task far more efficiently might be developed. Its motivations and preferences can be edited to be more consistent with each other. It will have an indefinite life span, be capable of reproducing, and transfer knowledge, skills, and code among its copies as well as cooperating and communicating with them better than humans do with each other.

The development of superintelligence from humans is another possibility, sometimes termed a weak superintelligence. It may come in the form of whole brain emulation, where a human brain is scanned and simulated on a computer. Many of the advantages a AGI has in developing superintelligence apply here as well. The development of Brain-computer interfaces may also lead to the creation of superintelligence. Biological enhancements such as genetic engineering and the use of nootropics could lead to superintelligence as well.