**Module 8**

**Beyond current generation AI and Toward Artificial General Intelligence**

Author: Sirwe Saeedi

**Outline**

**Current Artificial Intelligence**

**Artificial General Intelligence**

The aim of this module is creating a general snapshot of current advances in AI in a broad way and the built foundation of next generation of AI, lead to status of AGI and a speculative look at its future. This module is intended to give learners a general and partly comprehensive knowledge of strong AI.

**Current state of AI**

Most of us engage with AI daily and often do not even realize it. From search engine algorithms and customer support chatbots to enterprise apps and tools, AI is becoming a dependable and versatile technology. Amazon Alexa, Pandora and Netflix, YouTube, to name a few, all leverage AI technology to learn more about user profiles and preferences. Also, AI has become a technological reality for businesses and organizations across industries. AI has proven itself capable of improving process efficiency, reducing human errors and labor, and extracting insights from big data.

To better understand the current state of where AI can play a role, here are a few examples in business and many other people’ lives activities.

* Natural language processing for document review.
* Extraction of data from documents
* Detecting anomalies within data
* Learning from voice, image, and text data
* Automated customer support for online stores
* Personalized shopping experience including recommendations
* Forecasts and predictions for financial services
* AI curation and recommendations
* Face detection and security systems
* Internet of Things
* Siri, Apple’s voice assistant, uses AI to better understand your commands and habits and become more helpful over time. It may not yet be as receptive as a real assistant, but its accuracy has been improving incrementally.
* Gmail uses machine learning for security and privacy reasons—to stop spam and other unwanted emails.

… plus many other instances in every corner of our lives. Therefore, it might be clear to think about AI as a broad field, while Artificial General Intelligence (AGI) or strong AI is a more specific focus within it and what we have currently achieved is generally narrow and weak AI which has a narrow range of abilities and simulated intelligence [1],[2].  AGI is the concept of a machine with general intelligence that mimics human intelligence and/or behaviors, with the ability to learn and apply its intelligence to solve a range of problems. AGI can think, understand, and act in a way that is indistinguishable from a human in any given situation.

When AGI comes to a discussion, the main question of “what is the future of AI” comes along. Researchers have different opinions, and this make AGI as a grey area. Next section will give you the most popular definition and speculative look at AGI and its characteristics.

**What is AGI?**

The term “Artificial General Intelligence” (AGI) was coined by Mark Gubrud in 1997 then it was popularized by DeepMind’s chief scientist. The first book on current research on AGI, “Artificial General Intelligence” by Ben Goertzel, Cassio Pennachin [3] and annual AGI conference [4] have made AGI a common buzzword for superhuman AI. In fact, this book and AGI conference focus on engineering general intelligence, self-reflective, self-improving, and commonsensical intelligence. Therefore, three major concepts stand out as the key dimensions of AI: a fast self-improve, wishful thinking, and generalizing abilities of a human-like machine.

**Characteristics of AGI**

One of the AI pioneers John McCarthy wrote: “we cannot yet characterize in general what kinds of computational procedures we want to call intelligent”[[1]](#footnote-1), each of these dimensions is required to [5]:

* Automated reasoning that allows computers to reason completely under critical situation and uncertainty.
* Represent knowledge including commonsense reasoning of facts about the everyday world knowledge.
* Communicate in natural language with human and can understand the contents of documents.
* Learn from received data to make predictions or decisions.
* Act and respond to any observed situation.
* Creativity and empathy.
* Integrate all these skills toward common goals in AGI.

Many respected companies believe that we are nowhere close to making an AGI with multitasking abilities. However, companies like OpenAI, DeepMind, and Google Brain take achieving this goal very seriously [6]. That is why we cannot ignore enormous successes in AI technologies. Recent advances in AI include fake text, video, and audio generator and solve many problems only one at a time, kind of multi-tool machine to handle many different tasks without a new designed machine for each task, but these machines are far away from human-like capabilities.

### **Operational definitions of AGI**

There are several popular tests to consider AGI level achievement in increasing order as follow:

1. **The Turing Test**: (also known as the imitation game) is the most popular AI test proposed by Alan Turing back in 1950. It tests the ability of a computer, to observe whether it acts in an intelligent way or is distinguishable from human or not. Turing test has three parties, among which one is a computer, and two others are human. One of the two humans acts as a questioner and the other two parties are respondents. The evaluation of this test is for the questioner to decide whether the respondent is a machine or human. With slightly easy to answer questions, Turing test was able to pass, however, with more conversational questions that require a commonsense knowledge the test failed.
2. **The Coffee Test**: go into an average American house and figure out how to make coffee, including identifying the coffee machine, figuring out what the buttons do, finding the coffee in the cabinet, etc. Some AI researchers believe that AGI has arrived when a robot can perform the task of making coffee.
3. **The Robot College Student Test**:advocates an AI being enrolled in a college degree program and getting a degree using the same resources as other students enrolled for the same degree.
4. **The Employment Test**: One of the most common human activities that we go through is doing ‘jobs’, for which we are employed. In this test, the AI program should be able to perform jobs that are performed by humans. Then AGI can be evaluated on the basis of measuring the fraction of these jobs that are acceptably performed by the AI. The designer of the test said that AI programs must have at least the *potential* [to completely automate] economically important jobs.

Tasks for Learners (Open-ended Exploration)

**Investigate** operational tests of AGI like the ones highlighted above and the capabilities that true AGI will need to master. Learners can measure the progress toward human-level intelligence in other ways instead of trying operational tests. Similarly, numerous companies and research organizations are exploring alternative frameworks to measure progress based on granular human-equivalent capabilities, requirements to perform certain human tasks, or the combination of capabilities to perform every human job. Object detection, language understanding, social understanding capabilities of a child are human tasks that can be applied to measure AGI progress.

Many other examples of capabilities that learners can focus on to learn more about how AI can progress to AGI:

Sensory perception (determining spatial characteristic of an environment), natural language processing (AGI will need to be able to consume all sources of information with full comprehension), problem solving (In any general-purpose application, a robot (or an AI engine living in the cloud) will have to be able to diagnose problems, and then address them), creativity ( AGI will need rapidly to improve itself through a bootstrapping process to reach levels of intelligence), and social and emotional engagement (AGI system will need to understand humans, interpreting facial expressions or changes in tone that reveal underlying emotions).

Learners are recommended to explore the above further.

## **Can AI ever achieve general intelligence? How far are we from artificial general intelligence?** [7]

Although one or more of the AGI tests might seem compelling, the future of AGI is unclear. A self-driving car might indicate the arrival of AGI; however, it must act with high autonomy, at high speeds, in an extremely complex, dynamic, and uncertain environment. It must also face genuine moral dilemmas such as the philosopher’s trolley problem [8]. A Nature article [9], published in 2020, argues why AGI will not be realized, and it infers to one of the leading critics who says computers, who have no body, no childhood and no cultural practice, could not acquire intelligence at all.

However, today one might argue that new approaches to artificial intelligence research have made these arguments obsolete. Deep learning and Big Data are among the latest approaches, and advocates argue that they will be able to realize AGI eventually. A closer look reveals that although development of artificial intelligence for specific purposes has been impressive, we have not come much closer to developing AGI. The article further argues that this is in principle impossible.

This paper [10] attempts to prove that human is not equal to AGI and investigates that implicit assumption of equivalence between capabilities of AGI and human-level AI appears to be unjustified. “There may be such a thing as human-level AI, but human intelligence is nowhere near general”, Yann LeCun once said.

**Safety, Security, and Reliability (SSR) of future AGI**

Although, no proof exists that AGI will benefit humans and not harm or eliminate humans, it would be great to integrate SSR into technical developments toward achieving AGI, rather than having SSR thinking added as an afterthought. To ensure AGI technologies will not harm humanity and human/AGI values and goals are aligned, safeguards developed along SSR thinking should be in place. However, there are currently not many resources and studies available in the area of SSR development for AGI. This is currently an emerging area of research.

Tasks for Learners (Open-ended Exploration)

Investigate how SSR thinking can be incorporated into AGI development to safeguard future advances in AGI.

**Conclusion**

As no such human-level intelligence system has been demonstrated yet, AGI remains speculative. Some researchers believed that such intelligence is unlikely in the near future because it would require unforeseeable and fundamentally unpredictable breakthroughs. However, close attention needs to be given to the development path of AGI. Any AI system should be tested in transparent environment.

**References**

1. “Distinguishing between Narrow AI, General AI and Super AI”, <https://medium.com/mapping-out-2050/distinguishing-between-narrow-ai-general-ai-and-super-ai-a4bc44172e22>
2. “Types of AI: distinguishing between weak, strong, and super AI”, <https://www.thinkautomation.com/bots-and-ai/types-of-ai-distinguishing-between-weak-strong-and-super-ai/>
3. “Artificial General Intelligence” by Ben Goertzel,  Cassio Pennachin, <https://www.springer.com/gp/book/9783540237334#aboutBook>
4. Conference Series On AGI, <http://agi-conference.org/>
5. Areas Where AI Can Be Used, <https://medium.com/vsinghbisen/where-is-artificial-intelligence-used-areas-where-ai-can-be-used-14ba8c092e73>
6. “Artificial General Intelligence: Are we close, and does it even make sense to try?”<https://www.technologyreview.com/2020/10/15/1010461/artificial-general-intelligence-robots-ai-agi-deepmind-google-openai/>
7. “How Far Are We from Achieving Artificial General Intelligence?”, <https://www.forbes.com/sites/cognitiveworld/2019/06/10/how-far-are-we-from-achieving-artificial-general-intelligence/?sh=fcdd8926dc4d>
8. “Machine Ethics and the Trolley Problem”, <https://craigweich.com/post/36670778407/machine-ethics-and-the-trolley-problem-as>
9. “Why General Intelligence will not be realized”, Ragnar Fjelland, https://www.nature.com/articles/s41599-020-0494-4
10. “Human ≠ AGI”, Roman V. Yanpolskiy, <https://arxiv.org/pdf/2007.07710.pdf>

# General Reading:

1. [Russell, Stuart J.](https://en.wikipedia.org/wiki/Stuart_J._Russell); [Norvig, Peter](https://en.wikipedia.org/wiki/Peter_Norvig" \o "Philosophical Transactions of the Royal Society A) (2003), [Artificial Intelligence: A Modern Approach](http://aima.cs.berkeley.edu/)(2nd ed.), Upper Saddle River, New Jersey: Prentice Hall, [ISBN](https://en.wikipedia.org/wiki/ISBN_(identifier)) [0-13-790395-2](https://en.wikipedia.org/wiki/Special:BookSources/0-13-790395-2)
2. [Poole, David](https://en.wikipedia.org/w/index.php?title=David_Poole_(researcher)&action=edit&redlink=1); Mackworth, Alan; Goebel, Randy (1998), [Computational Intelligence: A Logical Approach](http://www.cs.ubc.ca/spider/poole/ci.html), New York: Oxford University Press, [archived](https://web.archive.org/web/20090725025030/http:/www.cs.ubc.ca/spider/poole/ci.html) from the original on 25 July 2009, retrieved 6 December 2007

1. For a discussion of some definitions of intelligence, see [philosophy of AI](https://en.wikipedia.org/wiki/Philosophy_of_artificial_intelligence) on Wikipedia [↑](#footnote-ref-1)