

# CS 1501 | Artificial General Intelligence | Spring 2020 | Syllabus

---

## Instructor

Mike Ferguson

## Email:

mef8dd@virginia.edu

## Faculty Sponsor:

Prof. N. Rich Nguyen

## Faculty Email:

nn4pj@virginia.edu

## Faculty Office Location:

Rice Hall, 204

## Meeting Time/Place:

Tuesdays 11:00-11:50

Online Synchronous

## Office Hours:

TBD

## Course Overview

The course serves to survey the exciting and nascent field of artificial general intelligence, or AGI. Throughout the semester, we will explore the intersection of computer science, philosophy, and cognitive science, as well as learn fundamental concepts of mind and superintelligence. The classes will mainly consist of looking at current research in the field, as well as debates and discussions about the nature of the mind, synthetic consciousness and intelligence, and what exactly it means to be *human*. Our generation will have to answer the most important questions humanity has ever asked regarding who we are, and we need to start early; we need to get this right.

The course is broken up into two main parts:

1. **Theoretical AGI** (*The Big Questions*): We will address, debate, and discuss the main questions in computer science of AGI, philosophy of mind, and cognitive neuroscience. See below to get an idea of some of the questions.
2. **Applied AGI**: We will look at the current research in the field, some challenges faced, and some paths to full-scale AGI. We will also explore the technical side of implementation and development of AGI, as well as research progress, programs, successes, and failures thus far.

I seek to make the class philosophical in nature while still firmly grounded in computer science. This field is unique, in the sense that it is strongly interdisciplinary, and I would encourage even those with no experience in cognitive science, computer science, or philosophy to take it. ***There are no prerequisites.*** This field, echoing what many of the top minds in it think, *requires* us to be interdisciplinary, and so my goal is to introduce philosophical and neuroscientific concepts to my fellow CS students to show how they connect, some of the gaps in those connections, and what that tells us.

*“The AI quest for artificial minds has transformed the mystery of consciousness into philosophy with a deadline.” – Dr. Max Tegmark, MIT Professor*

## Required Text

Nick Bostrom, *Superintelligence* (\$14.21 on Amazon)

Byron Reese, *The Fourth Age* (\$17.00 on Amazon)

## Course Schedule

Week	Lecture Material	Reading (Short!)
Week 1: <b>8/25/2020</b> (Introduction)	<u>What is AGI?</u> <ol style="list-style-type: none"> <li>1. Definition of AGI</li> <li>2. AGI vs AI vs machine learning</li> <li>3. The general intelligence problem</li> </ol>	<ol style="list-style-type: none"> <li>1. Read this: <a href="https://tinyurl.com/yxmvky nf">https://tinyurl.com/yxmvky nf</a></li> <li>2. No quiz due this week</li> </ol>
Week 2: <b>9/1/2020</b> (The Big Questions)	<u>The Cognitive Science of AGI</u> <ol style="list-style-type: none"> <li>1. Basic neuroanatomy</li> <li>2. Overview of cognitive functions</li> <li>3. Important concepts of mind</li> </ol>	<ol style="list-style-type: none"> <li>1. <i>The Fourth Age</i>, Chapter 13 and Chapter 14 (short)</li> <li>2. Quiz 1 due before class (on last week material)</li> </ol>
Week 3: <b>9/8/2020</b> (The Big Questions)	<u>Brief Philosophy of Mind</u> <ol style="list-style-type: none"> <li>1. The Turing Test</li> <li>2. Dualism vs Materialism</li> <li>3. The Chinese Room Argument</li> </ol>	<ol style="list-style-type: none"> <li>1. <i>The Fourth Age</i>, Chapter 5</li> <li>2. Quiz 2 due before class (on last week material)</li> </ol>
Week 4: <b>9/15/2020</b> (The Big Questions)	<u>Emotions, Feelings, and AGI</u> <ol style="list-style-type: none"> <li>1. What exactly are feelings and emotions?</li> <li>2. Could an AGI ever <i>feel</i> things, or <i>experience</i> emotions?</li> <li>3. Qualia and “What it’s like”</li> </ol>	<ol style="list-style-type: none"> <li>1. <i>The Fourth Age</i>, Chapter 16</li> <li>2. Quiz 3 due before class (on last week material)</li> </ol>
Week 5: <b>9/22/2020</b> (The Big Questions)	<u>Consciousness</u> <ol style="list-style-type: none"> <li>1. What is consciousness?</li> <li>2. Metacognition and selves</li> <li>3. Sentience</li> </ol>	<ol style="list-style-type: none"> <li>1. <i>The Fourth Age</i>, Chapter 18</li> <li>2. Quiz 4 due before class (on last week material)</li> </ol>
Week 6: <b>9/29/2020</b> (Practical AGI)	<u>Realistic Paths to Full Scale AGI - Part 1</u> <ol style="list-style-type: none"> <li>1. Cognitive Architectures</li> <li>2. LIDA and ACT-R</li> <li>3. SOAR and SIGMA</li> </ol>	<ol style="list-style-type: none"> <li>1. <i>Superintelligence</i>, Chapter 1</li> <li>2. Quiz 5 Due before class (on last week material)</li> </ol>
Week 7: <b>10/6/2020</b> (Practical AGI)	<u>Realistic Paths to Full Scale AGI- Part 2</u> <ol style="list-style-type: none"> <li>1. Seed AIs and recursive self-Improvement</li> <li>2. Gödel Machines</li> <li>4. Whole Brain Emulation</li> </ol>	<ol style="list-style-type: none"> <li>1. <i>Superintelligence</i>, Chapter 2</li> <li>3. Quiz 6 due before class (on last week material)</li> </ol>

Week	Lecture Material	Reading (Short!)
Week 8: <b>10/13/2020</b> (Practical AGI)	<u>Neuromorphic Computing and SNNs</u> <ol style="list-style-type: none"> <li>How biologically accurate are artificial neural networks today?</li> <li>Intro to Spiking Neural Networks (SNNs)</li> <li>Intel's Loihi Processor and IBM's True North Chip</li> </ol>	<ol style="list-style-type: none"> <li><i>Read This:</i> <a href="https://tinyurl.com/y4pctwlq">https://tinyurl.com/y4pctwlq</a></li> <li>Quiz 7 due before class (on last week material)</li> </ol>
Week 9: <b>10/20/2020</b> (Practical AGI)	<u>Superintelligence</u> <ol style="list-style-type: none"> <li>From AGI to superintelligence</li> <li>Establishment of a Singleton</li> <li>Kinetics of the intelligence explosion</li> </ol>	<ol style="list-style-type: none"> <li><i>Superintelligence</i>, Chapter 3</li> <li>Quiz 8 due before class (on last week material)</li> </ol>
Week 10: <b>10/27/2020</b> (Practical AGI)	<u>The infamous <i>Paper Clip Death Machine</i></u> <ol style="list-style-type: none"> <li>Terminal and instrumental goals</li> <li><i>Computronium</i> and instrumental convergence</li> <li>Orthogonality Thesis</li> </ol>	<ol style="list-style-type: none"> <li><i>Superintelligence</i>, Chapter 7</li> <li>Quiz 9 due before class (on last week material)</li> </ol>
Week 11: <b>11/3/2020</b> (Practical AGI)	<u>Concrete Problems in AI Safety</u> <ol style="list-style-type: none"> <li>Avoiding negative side effects</li> <li>Reward hacking</li> <li>Scalable oversight</li> </ol>	<ol style="list-style-type: none"> <li>No Reading. Watch this: <a href="https://youtu.be/AjyM-f8rDpg">https://youtu.be/AjyM-f8rDpg</a></li> <li>Quiz 10 due before class (on last week material)</li> </ol>
Week 12: <b>11/10/2020</b> (Practical AGI)	<u>The Impact of AGI</u> <ol style="list-style-type: none"> <li>Social impact</li> <li>Economic implications</li> <li>Political repercussions</li> </ol>	<ol style="list-style-type: none"> <li><i>Superintelligence</i>, Chapter 11</li> <li>Quiz 11 due before class (on last week material)</li> </ol>
Week 13: <b>11/17/2020</b> (Practical AGI)	<u>Conclusion</u> <ol style="list-style-type: none"> <li>What does it mean to be a machine? To be <i>human</i>?</li> <li>What will the world with full scale AGI look like?</li> <li>Will ego and human-chauvinism be our downfall?</li> </ol>	<ol style="list-style-type: none"> <li><i>Superintelligence</i>, Afterward (optional)</li> <li>Quiz 12 due before class (on last week material)</li> </ol>

Week	Lecture Material	Reading (Short!)
Week 14: 11/24/2020	<u>Class Celebration!</u> <ol style="list-style-type: none"> <li>1. Career/Research Paths</li> <li>2. Schools and Programs</li> <li>3. Extra Questions/Discussions</li> <li>4. Donuts Provided! (If class is in person)</li> </ol>	<ol style="list-style-type: none"> <li>1. No Reading</li> <li>2. Quiz 13 (Post Course Survey) due before class</li> </ol>

### Homework Policy

Since this is a class, we will have homework, but it will be brief. It will consist of short readings, all of which will be hand curated for maximal ratio of effort to take-away. I would also ask students to **do the reading before lecture, as to be somewhat familiar with the material before class**. This is also only a 1 credit class, therefore according to a rule of thumb, you should be spending around 3 hours/week on it. **The readings will not take that long**, and will be highly rewarding. Sometimes, there will only be a short YouTube video/article instead. You might say: “Gross, I took CS classes so I would not have to read!” To that I can relate; however, you are taking this class, so I would assume you like this stuff. As such, read the material and enjoy it; It will pay off!

### Grading Policy

The class is Pass/Fail. As such, as long as a majority (51/100) of points of the 100 possible points are earned, then you will pass. Attendance is the majority of the grade, with 3 excused absences allowed. Lastly, it is HIGHLY encouraged to debate and discuss in class, but not mandatory. I will keep track of those who like to discuss and seem motivated to learn what the class is about. ***If you like this stuff, come to class, and do the readings, then you should not have to worry about your grade.*** I do, however, have to officially evaluate learning somehow, and the best way is to administer quizzes on the readings / class material from the class before. They will be fairly simple, and if you do the reading and pay attention somewhat, then you will be fine. There is also a final paper/program on something relevant to the class, with details to come as the deadline approaches. The breakdown is as follows:

1. 50% - Reading Quizzes (Will drop lowest three: 13 total. Drop 3 = 10 quizzes graded, at 5 points each = 50 points)
2. 50% - Attendance (Three Excused Absences = same math as above = 50 points total)
3. Extra Credit (TBA), up to 5 additional points. Max grade in class = 105/100.

### Why any of this matters, or, *why you should take this class*

I’ve noticed from my personal experience that the best classes I have taken at UVA have had professors that put the material in perspective: that is, why what we are learning matters and why we should care. At the time of this writing, many top schools, such as MIT, Berkeley, Oxford, and others have or had recently offered classes on this subject (See MIT 6.S099: AGI for a phenomenal lecture series with the best in the field). Many of the field’s leaders, such as Elon Musk, Ray Kurzweil, Max Tegmark, Ben Goertzel, and Nick Bostrom have realized the urgency of the AGI and superintelligence problem, and how vastly unprepared humanity is for it. As stated above, we now live in a time where the biggest questions humanity has faced will not only have to be answered, but answered correctly, or the consequences could be catastrophic. We need as many bright minds working on this as possible, as we now enter in a new era, an era where humans could become gods or annihilate ourselves

spectacularly (See *Homo Deus*, by Yuvaal Noah Harari for another great read). The torch of intelligence and consciousness will pass from its former sole proprietor into the hands of a yet unknown entity, whose nature we cannot predict. If this does not terrify and excite you, look deeper. The world is changing; the future is now.

### About the Faculty Sponsor

This class would not be possible without the help and guidance of Professor N. Rich Nguyen, who teaches mainly CS 4774: Machine Learning, as well as (everyone's favorite CS class) CS 2150. He has given enormously good feedback on both educational topics and the curriculum, and any questions about the material can come his way as well as mine.

### About the Instructor

I (Mike) am a fourth-year computer science and cognitive science major, with a concentration in philosophy. My current research includes cognitive architecture implementation and design, neuromorphic computing, as well as convolutional neural network work with a graduate student, courtesy of UVA's Double Hoo Grant. I also spent the last summer at a remote internship for the University of Southern California doing AI/ML research. My hobbies include watches, stock (day/swing) trading, guitar, and the gym. I hope to research AGI/Superintelligence development and risks, and help ensure humanity doesn't accidentally kill itself. My eventual plan is to go to graduate school to get a Ph.D. in either computer science or cognitive neuroscience, and to become a professor; however, before that I am taking 5 years off to live in Sweden, learn the language, become an EU citizen and explore the world. *Om du talar svenska, vänligen kontakta mig och vi kan chatta!* Most importantly, I love playing fetch with my adopted Border Collie Shadow, the goodest of bois.

### Reviews From previous Semesters:

So you can get a taste of what to expect, I included some actual, verbatim reviews written by previous students:

1. I really loved everything that this course had to offer. While I was initially drawn into taking the class because of my passions in psychology and philosophy, the course opened my eyes to so much more. I went in knowing nearly nothing about the history and potential of Artificial Intelligence, let alone general intelligence, and came out of it knowing more than I thought was possible. The thought provoking questions and times for discussion were a great outlet for me because there hadn't been a lot of opportunities for that during my time in the E school. I am such a curious person and learning about something new and completely different is one of my favorite things so this class did a lot for me. Also the last lecture about meaning was so inspiring and exactly what I needed to hear in that moment. – *Anonymous, Spring 2020*
2. I very much enjoyed this course. I knew I was interested in AI as a concept, but going into this I didn't know what AGI was and now I realize that AGI is what really fascinates me most. Also, totally unexpected, I discovered an interest in philosophy. I didn't think I'd enjoy the more philosophical lectures, but they ended up being my favorite ones. I'm a first year, so I have plenty of time to explore things and now I know that I want a philosophy class to be part of that exploration. – *Anonymous, Spring 2020*
3. I loved it. It was a stress-free space where I could take the time to explore a topic I'm extremely interested in, and engage in fascinating conversations. I wish I had been able to do more of this sort of thing in my time at UVA. – *Anonymous, Spring 2020*
4. I really enjoyed taking this course, the lectures were always an interesting hour of my week and really made me think and question everything from our world to who we are and what makes us who we are. I liked that there was a lot of discussion

and questioning of our beliefs and every lecture had interesting, informative, and mind-boggling content. – *Anonymous, Spring 2020*

5. I really liked it, it was an awesome part of my week where I could go and get my mind blown and ponder my own existence lol. – *Anonymous, Spring 2020*
6. Really cool- helped me explore a field from an academic perspective and really dive deep and unpack these ideas. Now I can hold my own in a conversation and offer different perspectives on AGI. It's also helped me consider deeply what my purpose is and what humanity's purpose is... – *Anonymous, Spring 2020*
7. Great Class! one of my favorites. – *Anonymous, Spring 2020*
8. Great class. It's a pity that it was just a small 1 credit course. – *Anonymous, Spring 2020*
9. My favorite part of the course was talking about consciousness and what it means to be conscious. Literally made me question everything I knew and I got into several heated debates about it afterwards (very fun bc there isn't a right answer)! – *Anonymous, Spring 2020*
10. My favorite part was the quotes. As an engineering major whose favorite subject in school was always English, I really appreciated how you showed us the connection between literary and technical analysis. I looked forward to seeing how you'd tie each quote to your topic and I was never disappointed by how relevant the quote would turn out to be. – *Anonymous, Spring 2020*
11. I rate it 10/10 because it combines a lot interesting fields for an engaging discussion. – *Anonymous, Spring 2020*
12. I feel this course was very interesting and informative, I learned a lot of interesting things and engaged in great discussions. – *Anonymous, Spring 2020*
13. I really enjoyed it, and I thought it was a generally interesting and thought provoking course. – *Anonymous, Spring 2020*
14. Great class! Material was always interesting and class felt very engaging. – *Anonymous, Spring 2020*
15. I would say that this course is really enriching since it integrates knowledge in many fields – *Anonymous, Spring 2020*

**Further Readings/ Resources for those who are interested (extended list on Collab site):**

**1. Outstanding Books:**

- a. *Gödel, Escher, Bach: An Eternal Golden Braid* by Douglas Hofstadter
- b. *Artificial Minds* by Stan Franklin
- c. *The Consciousness Instinct* by Michael Gazzaniga
- d. *How the Mind Works* by Steven Pinker
- e. *A Brief Tour of Human Consciousness* by V.S. Ramachandran
- f. *Self Comes to Mind* by Antonio Damasio
- g. *Touching a Nerve: The Brain as Self* by Patricia Churchland
- h. *What to Think about Machines that Think* by John Brockman
- i. *From Bacteria to Bach and Back: The Evolution of Minds* by Daniel Dennet

**2. Best Videos:**

- a. MIT 6.S099: AGI Video Lecture Series on YouTube- Outstanding Interviews and Lectures by the very best of the field
- b. Lex Fridman's YouTube Channel – MIT Professor with Podcasts with, again, the best in the field
- c. Rob Miles YouTube Channel- Ph.D. Student at the University of Nottingham with great videos about the technical side of AGI
- d. Sam Harris's Ted Talk on AGI and losing control of it
- e. Anything by Christof Koch on consciousness

*“As our circle of knowledge expands, so does the circumference of darkness surrounding it.”  
- Albert Einstein*