**SHIT I KNOW**

**SHIT I KINDA KNOW**

**SHIT I DON’T KNOW**

**ANSWER**

Study Guide for the Final Exam ELEG 491 Spring 2020

All of the Required Readings, PowerPoints, and lectures from the course are potential subject matter for exam questions. To help you focus on the most salient features of that material, I’ve selected some topics, terms, and questions and listed them (by course topic/week) below.

1. ***Introduction: Ethical Concepts, Codes, and Theories***
2. **Philosophical and scientific definitions of ‘ethics’**
   1. Scientific
      1. an emergent behavior in social animals that constrains action by (primarily) non-violent sanction
   2. Philosophical
      1. a normative system of “oughts” to guide intentional actions that are preceded by choice
3. **Why engineering involves ethics**
   1. Ethics concerns the evaluation of action
      1. Designing and using technology is a form of (long lasting) action
         1. Therefore, part of ethics concerns the evaluation of technology
4. **Von Neumann’s (inevitability) thesis**
   1. An idea in the philosophy of technology stating that once a technology is introduced into a culture that what follows is inevitable development of that technology
      1. Cases of the first kind always lead to the second
5. **Definitions of Ethical Egoism, Utilitarianism, Kantianism, and Rights Views**
   1. Ethical Egoism:
      1. The theory that says you ought to pursue your own interests, and nothing else. It is influenced by the theory of games, and seeks rational strategies of cooperation in a “prisoner’s dilemma”
      2. Act in order to advance one’s own interests
   2. Utilitarianism:
      1. Utilitarianism is a family of consequentialist ethical theories that promotes actions that maximize happiness and well-being for the affected individuals
      2. Choose actions, the consequences of which are expected to maximize total social utility.
         1. maximizing social happiness
   3. Kantianism
      1. Act only on maxims (principles) that are always true
      2. [categorical imperative](https://en.wikipedia.org/wiki/Categorical_imperative)
         1. the criterion for whether a maxim is good or bad.
      3. Treat people as ends-in-themselves, never as mere means.
         1. all people are fundamentally rational beings
      4. this criterion amounts to a [thought experiment](https://en.wikipedia.org/wiki/Thought_experiment): to attempt to universalize the maxim (by imagining a world where all people necessarily acted in this way in the relevant circumstances) and then see if the maxim and its associated action would still be conceivable in such a world. For instance, holding the maxim *kill anyone who annoys you* and applying it universally would result in self termination. Thus holding this maxim is irrational as it ends up being impossible to hold it.
   4. Rights-based views
      1. agreed-upon lists of moral rights
      2. Rights-based theories give us “lists” of rights
6. **How the prisoner’s dilemma (pd) game relates to ethical egoism**
   1. The **prisoner's dilemma** is a standard example of a game analyzed in [game theory](https://en.wikipedia.org/wiki/Game_theory) that shows why two completely [rational](https://en.wikipedia.org/wiki/Rationality#Economics) individuals might not cooperate, even if it appears that it is in their best interests to do so

Two members of a criminal gang are arrested and imprisoned. Each prisoner is in solitary confinement with no means of communicating with the other. The prosecutors lack sufficient evidence to convict the pair on the principal charge, but they have enough to convict both on a lesser charge. Simultaneously, the prosecutors offer each prisoner a bargain. Each prisoner is given the opportunity either to betray the other by testifying that the other committed the crime, or to cooperate with the other by remaining silent.

* 1. The **ethical egoist** makes decisions on the basis of self interest. The basic presumption of the **Prisoner's Dilemma** is that the players are **egoists**.
  2. With two utilitarians in a **prisoner's dilemma**, both would rationally choose collaboration and both would be maximally satisfied with the outcome.

1. The outcome of a 2-person pd for a single play of the game
   1. If A and B each betray the other, each of them serves two years in prison
   2. If A betrays B but B remains silent, A will be set free and B will serve three years in prison
   3. If A remains silent but B betrays A, A will serve three years in prison and B will be set free
   4. If A and B both remain silent, both of them will serve only one year in prison (on the lesser charge).
2. **Why the (single-play) pd is a rational paradox**
   1. Because betraying a partner offers a greater reward than cooperating with them, all purely rational self-interested prisoners will betray the other, meaning the only possible outcome for two purely rational prisoners is for them to betray each other
3. **What Bentham’s hedonic calculus measures, and with what variables**
   1. Measures amount of pleasure an action will cause
      1. IDCFPE
      2. Intensity (I)
         1. How intense is the pleasure or pain?
      3. Duration (D)
         1. How long does the pleasure or pain last?
      4. Certainty ©
         1. What is the probability that the pleasure or pain will occur?
      5. Propinquity (nearness or remoteness) (N)
         1. How far off in the future is the pleasure or pain?
      6. Fecundity (F)
         1. What is the probability that the pleasure will lead to other pleasures?
      7. Purity (P)
         1. What is the probability that the pain will lead to other pains?
      8. Extent (E)
         1. How many persons are affected by the pleasure?
4. **The two versions of Kant’s Categorical Imperative**
   1. Formal Rule
      1. Act only on maxims (principles)that you could will universally;
   2. Material Rule
      1. Treat people as ends-in-themselves, never as mere means.
         1. Humans have an intrinsic value, and ought never to be valued like things that have merely instrumental value, e.g., tools
5. **Difference between negative and positive rights, with examples**
   1. Negative right
      1. FREEDOM FROM \_\_\_\_
      2. an enforceable claim such that, if person X possesses some good G, then all others must refrain from harming or diminishing (etc.) G.
         1. E.gr ight to live, to be free, freedom of speech, freedom of religion, freedom from violence, freedom from slavery, and property rights.
      3. If the right is infringed, X must be “made whole” again.
   2. Positive right
      1. FREEDOM TO \_\_\_\_
      2. an enforceable claim such that X must be given (some amount of) G, after which point X also gains a negative right to G.
         1. E.g rights to free schooling, free healthcare, a job, and a minimum wage.
6. ***Privacy***
7. **3 possible meanings of a right to privacy (relative to kinds of freedom)**
   1. Freedom from unwarranted intrusion
   2. Freedom from interference in personal affairs or personal decisions
   3. Freedom to control the flow of personal information.
8. **McNealy's views of privacy**
   1. (1999) Consumer privacy issues are a “red herring,” McNealy told a group of reporters that year. “You have zero privacy anyway. Get over it.”
   2. (2015) “It doesn’t really bother me that Google and AT&T have information about me, because I can always switch to another provider... If Uber starts screwing around with my data, I’ll use Lyft…. [but] It scares me to death when the NSA or the IRS know things about my personal life and how I vote.”
9. **Definition of privacy from Warren and Brandeis article**
   1. being let alone, (Warren & Brandeis, 1890)
   2. being free from government intrusion (e.g., into one’s physical space)
10. **What the 4th amendment to the U.S. Constitution protects**
    1. Protection against unwarranted search and seizures
    2. In addition, it sets requirements for issuing [warrants](https://en.wikipedia.org/wiki/Warrant_(law))
       1. warrants must be issued by a judge or magistrate, justified by [probable cause](https://en.wikipedia.org/wiki/Probable_cause), supported by oath or affirmation, and must particularly describe the place to be searched and the persons or things to be seized
11. **Interference/intrusion, tort, reductionist, and epistemic views of privacy**
    1. Interference/intrusion
       1. being let alone, (Warren & Brandeis, 1890)
       2. being free from government intrusion (e.g., into one’s physical space)
    2. Tort
       1. Appropriateness of others getting info. @ you
       2. Technology governing access/distribution
       3. Laws (statutory and common law) about injury
    3. Van den Hoven on “Reductionist” views of privacy
       1. It is not privacy we’re trying to protect, but “property rights, security, autonomy, intimacy or friendship, democracy, liberty, dignity, or utility and economic value.”
    4. Epistemic (relational) view of privacy
       1. “Having privacy means that others don't know certain private propositions.” Cf. Bentham’s “Panopticon”
12. **What was Bentham’s panopticon?**
    1. The concept of the design is to allow all prisoners of an institution to be observed by a single [security guard](https://en.wikipedia.org/wiki/Security_guard), without the inmates being able to tell whether they are being watched.
       1. Although it is physically impossible for the single guard to observe all the inmates' cells at once, the fact that the inmates cannot know when they are being watched means that they are motivated to act as though they are being watched at all times.
13. **Van den Hoven’s 4 moral reasons for protecting personal data**
    1. Prevention of harm
       1. through impersonation, identity theft, etc.
    2. Informational inequality
       1. commodification of personal information
    3. Informational injustice and discrimination
       1. health care and insurance, contracts, employment, etc.
    4. Encroachment on moral autonomy
       1. outside influence
14. **Examples of technologies that threaten privacy** 
    1. Sensors on the nanoscale
    2. Ambient IT
    3. Genome sequencing
    4. Facial recognition software
15. **Examples of technological solutions to privacy problems**
    1. Value (Privacy) Sensitive Design
    2. Privacy Enhancing Technologies
       1. (Tor and Proxy Servers, Freenet)
    3. Cryptography, Identity Management
16. **Examples of personal data**
    1. Trivial
17. **Comparison of EU and US views of protecting data (Van den Hoven video)**
    1. Data protection is a European framework
       1. Put a fence around vulnerable data
       2. Constrain access to the things you think are vulnerable
    2. Data is a use in the US context
       1. It's a commodity
18. ***Intellectual Property***
19. **Locke's account of property**
    1. A claim on the fruits of labor on things taken from “the commons”
    2. It is very clear, that God … has given the earth to the children of men [and] to mankind in common. But this being supposed, it seems to some a very great difficulty, how any one should ever come to have a property in any thing. But I shall endeavour to shew, how men might come to have a property in several parts of that which God gave to mankind in common, and that without any express compact of all the commoners…. The earth, and all that is therein, is given to men for the support and comfort of their being. And tho' all the fruits it naturally produces, and beasts it feeds, belong to mankind in common
20. **“Economics of law” view of property**
    1. From an “economics of law” view, we want the optimal mix of protection of “fences” and “codes”
21. **4 schemes of intellectual property (IP) protections & their durations**
    1. Copyrights (life of author + 70 years)
    2. Patents (20 years past filing date)
    3. Trademarks (length of use)
    4. Trade secrets (length of secrecy)
22. **What is being “balanced” in Article 1 Section 8 of U.S. Constitution**
    1. (Article 1, Section 8) states:
       1. The congress shall have the power... to promote the Progress of Science and the useful Arts, by securing for limited Times to authors and inventors the exclusive Rights to their respective Writings and Discoveries.”
23. **Lessig’s 4 modes of regulation (give examples of each)**
    1. Norms
    2. Markets
    3. Codes (real and software)
    4. Law (only necessary when others fail)
24. **What the Bayh-Dole Act (as amended) does**
    1. Establishes IP rules between Fed. Gov’t and “contractors”
    2. Allows universities to own IP from research
    3. Allows Gov’t to use IP that they fund free of charge
    4. Requires scientists to claim title to discoveries or risk losing that property
25. **Tangible vs. Intellectual objects: compare and contrast**
    1. Tangible objects are excludable and rivalrous in nature. (If I own an apple, you can’t own it, and thus I can “exclude” your access to it. If I eat it, you can’t also eat it—there is less apple to go around)
    2. Intellectual objects, such as software programs, are non-excludable.
    3. Scarcity applies to tangible objects (in markets this gives rise to competition and rivalry)
    4. Scarcity need not exist for intellectual objects; intellectual objects can be easily reproduced.
26. **Exclusion and Rivalry (in economics)**
    1. In economics, a good is said to be rivalrous or a rival if its consumption by one consumer prevents simultaneous consumption by other consumers, or if consumption by one party reduces the ability of another party to consume it.
    2. the exclusion principle states "the owner of a private good may exclude others from use unless they pay.";
27. **How ideas get “expressed” or “fixed” and copyright protected**
    1. If an idea is literary or artistic in nature, it must be expressed (or "fixed") in some tangible medium in order to be protected. This is copyright.
    2. Functional ideas, such as an inventions, must be expressed in terms of a machine or a process. This is patent.
28. Jefferson's argument against inventions as property
    1. “If nature has made any one thing less susceptible than all others of exclusive property, it is the action of the thinking power called an idea, which an individual may exclusively possess as long as he keeps it to himself; but the moment it is divulged, it forces itself into the possession of every one, and the receiver cannot dispossess himself of it. **Its peculiar character, too, is that no one possesses the less, because every other possess the whole of it. He who receives an idea from me, receives instruction himself without lessening mine**…
    2. That ideas should freely spread from one to another over the globe, for the moral and mutual instruction of man, and improvement of his condition, seems to have been peculiarly and benevolently designed by nature, when she made them, like fire, expansible over all space, without lessening their density at any point... Inventions then cannot, in nature, be a subject of property.”
       1. Thomas Jefferson, Letter to Isaac McPherson, Monticello, August 13, 1813
29. **Complications in considering software as IP**
    1. Should computer programs be eligible for patent protection?
    2. How did they get protected by copyright law?
    3. Do they deserve both, or perhaps neither, kind of protection?
    4. Computer software consist of lines of programming code (or codified thought).
    5. It is not expressed or "fixed" in a tangible medium in the same way that literary works are.
    6. Software programs also resemble algorithms, which, like mathematical ideas or "mental steps," are not eligible for patent protection.
    7. Eventually, however, both copyright and patent protections were granted to software programs.
30. **Exclusive rights of copyright holder (what copyright protects)**
    1. Copyright holders have the exclusive right to:
       1. make copies of the work;
       2. produce derivative works, translations into other languages, movies based on the book, and so forth;
       3. distribute copies;
       4. perform works in public (musicals, plays. etc.);
       5. display works in public (e.g., art works).
       6. Fair use and First Sale doctrines apply
31. **Importance of *Sony v. Universal Studios,* and *White-Smith v. Apollo* cases**
    1. *Sony v. Universal Studios*
       1. *ruled that the making of individual copies of complete* [*television*](https://en.wikipedia.org/wiki/Television) *shows for purposes of* [*time shifting*](https://en.wikipedia.org/wiki/Time_shifting) *does not constitute* [*copyright infringement*](https://en.wikipedia.org/wiki/Copyright_infringement)*, but is* [*fair use*](https://en.wikipedia.org/wiki/Fair_use)
    2. *White-Smith v. Apollo*
       1. *ruled that manufacturers of music rolls for* [*player pianos*](https://en.wikipedia.org/wiki/Player_piano) *did not have to pay* [*royalties*](https://en.wikipedia.org/wiki/Royalties) *to the composers*
       2. *ruling was based on a holding that the piano rolls were not copies of the plaintiffs' copyrighted sheet music, but were instead parts of the machine that reproduced the music*
32. **Importance of DMCA for copyright protection: anti-circumvention**
    1. FUCK DMCA FUCK DENUVO
33. **What is *fair use***
    1. *Fair use balances the exclusive controls given to copyright holders against the broader interests of society (remember Art. 1 Sec.8)*
    2. *Fair use means that someone may make limited use of another person's copyrighted work for purposes such as:*
       1. *criticism,*
       2. *comment,*
       3. *teaching,*
       4. *scholarship,*
       5. *research.*
34. What is *first sale* doctrine
    1. provides that an individual who knowingly purchases a copy of a copyrighted work from the copyright holder receives the right to sell, display or otherwise dispose of that particular copy, notwithstanding the interests of the copyright owner
35. **How specifically did Napster contribute to copyright infringement?**
    1. Napster (1999) was sued by the Recording Industry Association of America (RIAA) for distributing copyrighted music on the Internet through a centralized server.
36. **Conditions for gaining a patent**
    1. Patents are granted to inventions and discoveries that satisfy three conditions:
       1. usefulness,
       2. novelty,
       3. non-obviousness.
37. **First-to-file vs. First-to-invent**
    1. In a first-to-file system, the right to the grant of a patent for a given invention lies with the first person to file a patent application for protection of that invention, regardless of the date of actual invention
    2. **First-to-Invent**" doctrine, which means that the inventor who **first** conceived of the invention and then diligently reduced it to practice by filing a patent application (or actual reduction to practice) is considered the **first** inventor and is entitled to patent
38. **Significance of *Gottschalk v. Benson (1972)***
    1. *ruled that a process claim directed to a numerical algorithm, as such, was not patentable because "the patent would wholly pre-empt the mathematical formula and in practical effect would be a patent on the algorithm itself."*
39. **Different Creative Commons licenses**
    1. CC expands the range of creative work available to others legally to build upon and share.
    2. The CC menu provides four options:
       1. Attribution—Permit others to copy, distribute, display, and perform the work and derivative works based upon it only if they give you credit;
       2. Noncommercial—Permit others to copy, distribute, display, and perform the work and derivative works based upon it only for noncommercial purposes;
       3. Derivative works—Permit others to copy, distribute, display, and perform only verbatim copies of the work, not derivative works based upon it;
       4. Share alike—Permit others to distribute derivative works only under a license identical to the license that governs your work.
40. **Positions/arguments of the FSF on the "balancing view" of IP**
    1. Challenges to the “balancing” view: Free Software Foundation (FSF)
    2. run the program, for any purpose;
    3. study how the program works, and adapt it for your needs;
    4. redistribute copies so you can help your neighbor;
    5. improve the program, and release your improvements to the public so that the whole community benefits.
41. ***Security and Safety***
42. **Main difference between security & safety, according to Leveson & Young**
    1. Safety
       1. prevent losses due to unintentional actions by benevolent actors
    2. Security
       1. prevent losses due to intentional actions by malevolent actors”
43. **Features of the Failure (or Event-Chain) accident model**
    1. Isolate the proximate or root cause
    2. Label all other conditions contributory
    3. Focus on direct causality (necessary conditions, i.e. if event X had not occurred, then the following event Y would not have occurred) and ignore reasons for failure
    4. This might have worked for simpler electro-mechanical systems, but doesn’t work for many safety-critical systems
44. **Difference between necessary conditions and sufficient conditions**
    1. A **sufficient condition** guarantees the truth of another **condition**, but is not **necessary** for that other **condition** to happen.
    2. A **necessary condition** is required for something else to happen, but it does not guarantee that the something else happens
45. **Why accident analysis is interested in *reasons***
46. **Why Failure model doesn’t work for safety-critical computer systems**
47. **Two kinds of failure (Leveson); how Mars Polar lander exemplifies one**
    1. Individual Component(s) failure
    2. Dysfunctional interactions btw. non-failed components (often unpredictable)
    3. E.g. the Mars Polar Lander descent control SW
       1. “In this and in most software-related accidents, the software operates exactly as specified, that is, the software, following its requirements, commands component behavior that violates system safety constraints or the software design contributes to unsafe behavior by human operators.”
48. **Definition of an emergent property**
    1. emergence occurs when an entity is observed to have properties its parts do not have on their own.
    2. These properties or behaviors emerge only when the parts interact in a wider whole.
       1. For example, smooth forward motion emerges when a bicycle and its rider interoperate, but neither part can produce the behavior on their own.
49. **Safety as an emergent property**
    1. **safety** is an **emergent property**" actually means the opposite, namely that the lack of **safety** - accidents, incidents, and other forms of adverse outcomes - is **emergent**
50. **Main features of Leveson’s STAMP model**
    1. **STAMP (System-Theoretic Accident Model and Processes) is the name of the new accident causality model based on systems theory**
       1. Hierarchy of levels of organization
       2. Safety as an emergent property at a level
       3. (E.g. questions about the safety of a single valve in a plant are meaningless; we have to know how the value is used and controlled)
51. **Role of constraints in STAMP**
    1. Accidents arise from inadequate control or enforcement of safety-related constraints on the design, development, and operation of the system.
52. **Cantwell-Smith’s account models, and how computers depend on them**
    1. Note computers’ special dependence on models:
       1. “you write an explicit description of the model down inside the computer, in the form of a set of rules or what are called representations… the facts and data thought to be relevant to the system’s behavior.”
53. **What is formal program verification? (C-S)**
    1. involves proving that a **program** satisfies a **formal** specification of its behavior
54. **Left- and Right-hand relationships in <Computer-Model-World> complex**
55. **Why is “correctness” not sufficient for safety, according to C-S?**
    1. The only way to prove the **correctness of an algorithm** over all possible inputs is by reasoning formally or mathematically about it
56. **Arguments for/against hacking**
    1. Trivial
57. **Details of two hacker "ethics" (MIT and Levy)**
    1. MIT
       1. A hack must:
          1. be safe
          2. not damage anything
          3. not damage anyone, either physically, mentally or emotionally
          4. be funny, at least to most of the people who experience it
    2. Levy
       1. Access to computers should be unlimited and total.
       2. 2. Always yield to the Hands-On Imperative
       3. 3. All information should be free.
       4. 4. Mistrust authority--promote decentralization.
       5. 5. Hackers should be judged by their hacking.
       6. 6. You can create art and beauty on a computer.
       7. 7. Computers can change your life for the better.

***V. Engineering Failures***

* 1. **Identify and explain the 3 approaches to engineering failures that we read**
     1. Monteiro- an (idealistic) Designer’s Code of Ethics: (designer is responsible for purpose as well as form)
     2. Brey- an “anticipatory” ethics of forecasting (designer must foresee ethical issues & correct)
     3. Regalado & Blyler- a retrospective “worst of” case-studies approach
  2. **Where (primarily) do ethical filters appear in science and engineering?**
     1. between the engineering and commercial distribution phases
  3. **What else functions as a filter to stop some technologies from reaching the marketplace?**
     1. Regulations, tax incentives, patent laws, personal ethical views, environmental considerations...ect
  4. **Oppenheimer's approach to the problem of possibly evil technologies**
     1. that as scientists we must continue to discover more about the world and have faith that the power we give the world won’t be misused
  5. **Brey's 3 levels of "anticipatory" ethics**
     1. Technology level: a collection of related techniques with a common domain/purpose
     2. Artifact Level: a thing that when used in a proper manner performs a desired result
     3. Application Level: using the artifact to do the thing
  6. **Brey's definitions of *technology, technique, artifact,* and *application***
     1. *Technique: a procedure to accomplish a speciﬁc activity or task.*
     2. *others shown above*
  7. **Contents of Brey's "checklist" for anticipatory ethics**
     1. Harms and risks
     2. Rights
     3. Autonomy
     4. Human Dignity
     5. Privacy
     6. Property
     7. Justice
     8. Well-being and the Common Good
  8. **Bill Joy’s argument about “why the future doesn’t need us”**
     1. Technology is advancing to the point where very powerful things like nano technology and genetic engineering will be available to us or at least small groups of people cheaply.
  9. **Stuart Russell's comparison of AI and nuclear technologies**
     1. both seem great until we consider the risks, neither had enough regulation upon their entering the world
  10. **Explain the case of nanotechnology development and its ethical implications**
      1. can provide a lot of good. nanoparticles have novel properties.
      2. not much is yet known about the effects of nano technologies. ie Toxicity, effect on environment, possibility for use in spying.
      3. causes concern about future health and privacy
  11. **Describe the "biggest technology failures" of 2018-19**
      1. boeing big oof’d, chinese creeper guy, facebook sucks
  12. **Be able to explain the main tenets of Monteiro's "Designer Code of Ethics"**
      1. He is a pretentious man who preaches a lot about how “a broken gun is better designed than a working one”.
      2. He believes that designers have a moral responsibility to only design things that are good for the world.

***VI. Autonomous Weapons & Cyberwar***

* 1. **Definition of automation (NASA 1989**
  2. the explicit allocation of functions between humans and machines
  3. **Differences: Automation vs. Autonomy**
  4. Automation = human involved in the process
  5. Autonomy = computer decides what it wants to do and how it wants to do it all on its own
  6. **Ethical importance of autonomy**
  7. If a machine is able to think for itself does it deserve rights
  8. **Principles of Just War theory (*ad bellum* and *in bello*) and Laws of War**
  9. The “right to go to war” (jus ad bellum): right purpose, due authority, last resort, proportionality
  10. The “right conduct in war” (jus in bello): noncombatant immunity, proportionality
  11. that have been integrated with the Laws of War
  12. One ethical focus of most versions: minimizing casualties/proportionality
  13. **Difference between micro-proportionality and macro-proportionality**
  14. macro = goal of the war overall, and the reason for its starting
  15. micro = does a specific operation/action in a war and its effect justify its cost
  16. **Arkin's reasons that autonomous robots will perform better than humans**
  17. can act conservatively since don’t need to protect themselves
  18. can have sensors better than humans for selecting correct targets
  19. no emotions to cloud judgment
  20. monitor ethical behaviour of human soldiers
  21. **Main arguments against autonomous warfare robots**
  22. could fall into the wrong hands
  23. dehumanize war
  24. terrorize the populace
  25. do they respect the populace
  26. **Relation of hacking to the “technical revolution” according to Dr. Crash**
  27. computer had been in the hands of big government and business until that point. hacking is a way to fight computer monopolies
  28. **4 arguments for teaching offensive hacking techniques**
  29. The “good guys have to keep up” argument
  30. The “I can trust my students” argument
  31. The “we aren’t selling exploits” argument
  32. The “don’t drive research underground by regulation” argument
  33. **Origin of the idea of war “within limits”**
  34. ww1 gas warfare was so bad people decided to not do that anymore.

***VII.******E****-****Commerce and Free Speech Online***

* 1. **Two Senses of "Internet Regulation" (give examples of both)**
  2. regulating content, e.g., online pornography, hate speech, etc.
  3. regulating processes – i.e., rules and policies – for commercial transactions in cyberspace.
  4. **Alternatives for who/what should regulate commerce and speech**
  5. the Federal/state government
  6. private organizations
  7. Internet users themselves
  8. **Four traditional types ("title") of media**
  9. publishing
  10. broadcast
  11. distribution
  12. common carrier
  13. **Current FCC/Net Neutrality dispute over Internet and common carriage**
  14. FCC got rid of net neutrality, some concerned that this will alow companies to discriminate on what sort of traffic they allow, and how much they charge for it
  15. The FCC wont let me be or let me be me so let me see, they try to shut me down on MTV but it feels so empty without me
  16. **Fate of Bell/AT&T as a common carrier**
  17. Bell’s patents expired by 1894
  18. Bell consolidated control over long-distance service nationwide by the 1920’s
  19. The 1934 Act declares Bell a “natural monopoly” and places it under common carrier regulation—what they “transported” had to treated with neutrality
  20. AT&T converges with the “baby Bells to continue the monopoly
  21. **Wu’s principle of net neutrality and comparison to other media**
  22. “a maximally useful public information network aspires to treat all content, sites, and platforms equally.”
  23. **e. Examples of content- and process-based federal regulatory agencies**
  24. Examples of content-based agencies include:
      1. Food and Drug Administration (FDA);
      2. Local and State Boards of Health
      3. ;Liquor Control Board.
  25. Examples of process-based agencies include:
      1. Federal Trade Commission (FTC);
      2. Federal Communications Commission (FCC)
      3. Security Exchange Commission (SEC).
  26. **What does DRM allow in cyberspace?**
  27. DRM technologies allow content owners to regulate the flow of information in digital media by blocking access to it via encryption, and limiting access through passwords.
  28. **History of Internet Domain Names**
  29. Originally, one person (Jon Postel), administered the Internet registries. The National Science Foundation (NSF) formerly controlled the licensing of domain names in the 1980s (taking control from DARPA)
  30. ICANN (Internet Corporation of Assigned Names and Numbers) took over the process in 1998; it was seen as more business friendly than NSF
  31. Now an international board (Governmental Advisory Committee) administers DNS and other protocols.
  32. **First Amendment of the U.S. Constitution & how it is a conditional right**
  33. According to the First Amendment to the U.S. Constitution, “Congress shall make no law...abridging the freedom of speech, or of the press...”However, the right to free speech is a conditional right (not an absolute right).
  34. **Does the First Amendment restrict hate speech?**
  35. Not really but you can’t make direct threats against people

l. Obscenity standards in *Miller v. California*

1. *The Miller v. California case (1973) established that material is obscene when it:1.depicts sexual (or excretory) acts whose depiction is specifically prohibited by law. 2.depicts these acts in a patently offensive manner, appealing to prurient interest as judged by a reasonable person using community standards. 3.has no serious social, literary, artistic, political, or scientific value. (The SLAPS test)Indecent speech can be restricted in terms of time, manner, and place.*
2. m. Fate of speech-limiting legislation: CDA, COPA, and CIPA
3. n. Mill’s “Harm Principle” and its relation to the hate speech debate

***VIII. Online Community and Social Media***

**a. Standard and “online” sociology of community**

1. Standard definitions usually refer to people living in the same place (town, city, etc.) and sharing (mostly) the same norms and laws. With the Internet, our understanding of community has expanded to include online interactions, e.g., chat rooms, forums, games, and social networking sites such as Facebook.

**b. Borgman, Dreyfus, and Contemporary critiques of online communities**

1. Borgmann: online communities compete with RL and cause disconnectedness and “diminished presence”
2. Dreyfus: anonymity online is without risk, hence one can’t build real relationships (even education!)
3. Contemporary critiques: online communities might support polarization, degenerate behavior, and in some cases criminality (Anonymous)

**c. Relation of Plato’s Ring of Gyges to online behavior**

1. Gyges gets a ring that makes him invisible, he does lots of bad stuff. Does anonymity make people do bad things

**c. Details of LamdaMOO case**

1. dude virtualy rapes a girl in an online comunity
2. Rape in Cyberspace" describes a "cyberrape" in a multi-player computer game or [MUD](https://en.wikipedia.org/wiki/MUD) called [LambdaMOO](https://en.wikipedia.org/wiki/LambdaMOO) that took place on a Monday night in March of 1993 and discusses the repercussions of this act on the [virtual community](https://en.wikipedia.org/wiki/Virtual_community) and subsequent changes to the design of the MUD program
3. LambdaMOO's main creator set up a system of [petitions](https://en.wikipedia.org/wiki/Petitions) where anyone could put to popular vote anything requiring administrative powers for its implementation.
   1. Through this system, LambdaMOO users put into place a @boot command, which temporarily disconnects disruptive guest users from the server.

**d. Philosophical claims about the wrong done in the LamdaMOO case**

1. Online speech is a form of action (Speech Act Theory) and thus ”real”
2. Identification or attachment between avatar and controller is key to understanding the wrong(s). So are reasonable expectations.
3. Whether text-based or graphical (3D) depictions matter to virtual morality is controversial (cf. Wolfendale)

***IX.Technology and Democracy***

**a. Types of democracy along the "continuum"**

1. A “continuum” concept: individual choice in government decisions
2. Ancient Greece: The demos were the commoners—not necessarily “majority rule”
3. Direct Democracy <> Plebiscitarian
4. Representative Democracy <> Republican

**b. Relevance of Arrow's Theorem to the issue of social choice**

1. In short, the theorem states that no rank-order electoral system can be designed that always satisfies these three "fairness" criteria:
   1. If every voter prefers alternative X over alternative Y, then the group prefers X over Y.
   2. If every voter's preference between X and Y remains unchanged, then the group's preference between X and Y will also remain unchanged (even if voters' preferences between other pairs like X and Z, Y and Z, or Z and W change).
   3. There is no "[dictator](https://en.wikipedia.org/wiki/Dictator)": no single voter possesses the power to always determine the group's preference.
2. There is no one “reasonable” procedure to derive social choice from individual choice

**c. Example of non-transitivity problem (Condorcet) in ordered-preference voting**

1. if all the voters vote for different things in a different order. no transitive solution since don’t prefer any one choice over another reliably
2. look at the table example

**d. Definition and examples of the "digital divide"**

1. The “digital divide” separates the information “haves” and “have-nots”
2. poor countries might not have computers

**e. General IT trends in lessening the global digital divides**

1. used to be that almost all internet users were in western countries, now it is more common in developing nations

**f. Other IT-related political problems for democracy**

1. Auto-indoctrination
2. Personalization filters (echo chamber effect)
3. Mass dispersion of half-truths (and lies)
4. Commenting/Trolling
5. Information overload•Data mining to make the voter=“the product”

**g. Two ways that IT has impacted elections**

1. Political blogs spread false info
2. propaganda spreads quickly

**h. Carother's main reasons that IT hasn't helped democracy**

1.It is too soon to tell—some changes may be “local”

2.Large, countervailing trends (e.g. authoritarianism)

3.IT itself doesn’t mobilize “collective action”

**i. Examples (from Carothers) of IT/democracy successes and failures**

1. shrewd authoritarian leaders are able to limit the empowering effects of technology on ordinary citizens both by cracking down on civic space — which they are doing in many countries — and by using new technologies for their own anti-democratic ends, such as tracking the whereabouts and actions of pro-democracy activists.

**j. Lessons from the use of electronic devices in the "Arab Spring"**

1. Protestors threatened to bring down the government, so the Mubarak administration shut down the country’s Internet services and mobile phone resources.
2. But the protestors had already planned demonstrations via Facebook and Twitter before the online services in Egypt could be shut down.

**k. Marcuse's argument about freedom in advanced industrial civilization**

1. A comfortable, smooth, reasonable, democratic unfreedom prevails in advanced industrial civilization
2. Rational “industrial” development can both supply and deprive freedom
3. Technological progress can make us (voluntarily) part of ”the machine”
4. The “machine process” integrates all aspects of work, culture, and leisure
5. Potentially, automation can satisfy all our “needs” and make us “free”

**m. Marcuse's kinds of freedom**

1. Freedom of thought, speech, and conscience = old freedoms
2. Freedom from want = concrete substance of all freedom
3. need new freedoms

**n. Marcuse's concept of the "machine process"**

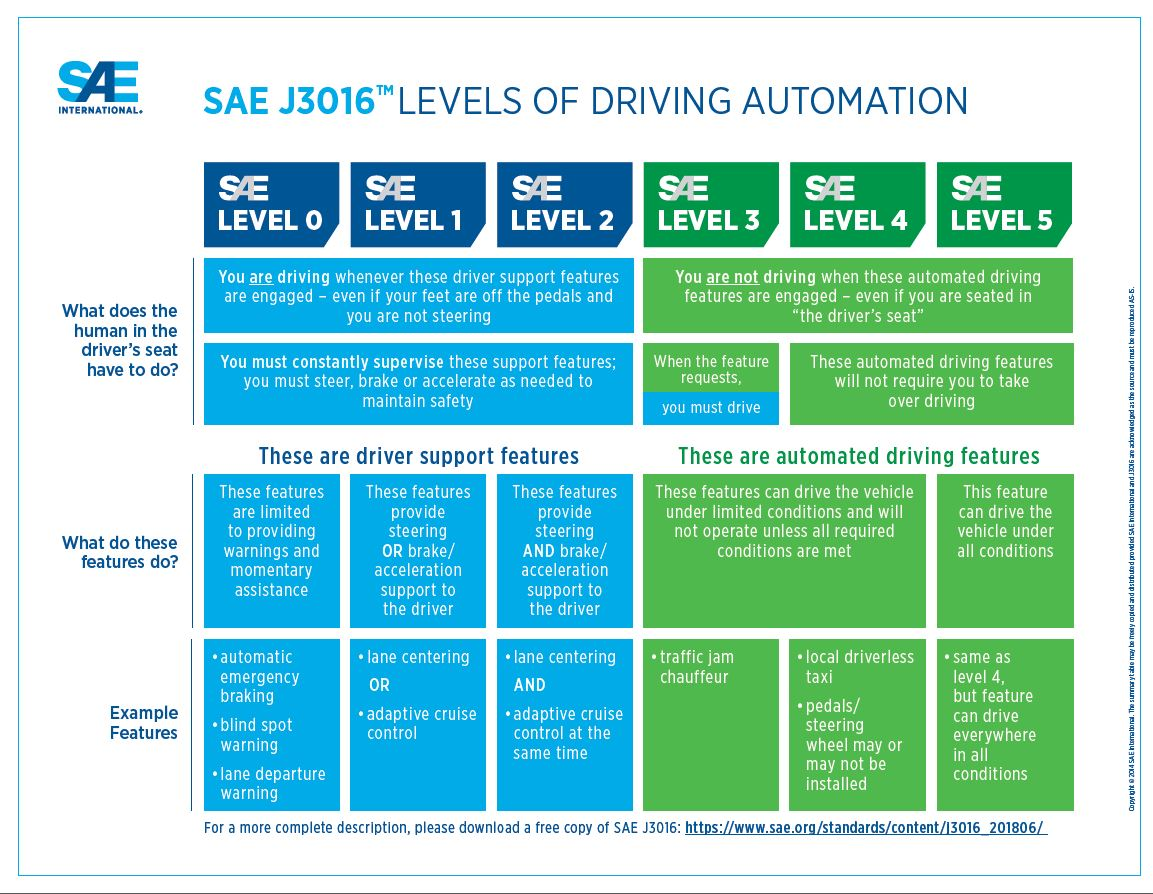
1. Today political power asserts itself through its power over the machine process and over the technical organization of the apparatus.
2. The government of advanced and advancing industrial societies can maintain and secure itself only when it succeeds in mobilizing, organizing, and exploiting the technical, scientific, and mechanical productivity available to industrial civilization. And this productivity mobilizes society as a whole, above and beyond any particular individual or group interests

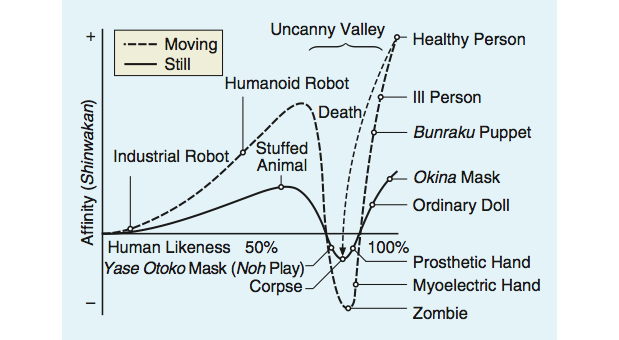
**o. Marcuse's argument future progress an automation**

1. not having to do any work may allow us to focus on other things and gain new freedoms. universal basic income and automation may allow people more freedom.

***X. Robotics and Autonomous Systems***

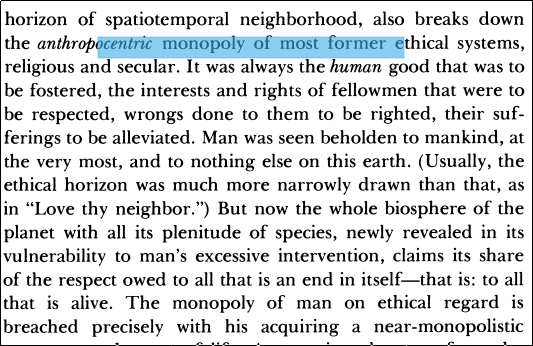
* 1. **What are Lin's examples of AI decisions that fall into a "moral gray space"?**
     1. there are also decisions that are neither obviously right nor wrong, as they fall within the shadowy but sizable space of judgment calls. Those judgment calls, crystallized as code, are the ones that demand serious ethical consideration, especially since they may raise challenges to risk and liability.
  2. **Themes in the ethics of automation as researched by different disciplines**
     1. Risk
        1. engineering problem w/values
     2. Autonomy
        1. philosophy
     3. Interaction, Trust and Acceptance
        1. psychology
  3. **How is the traditional Trolley Problem supposed to present a moral dilemma?**
     1. There is a runaway trolley barreling down the railway tracks. Ahead, on the tracks, there are five people tied up and unable to move. The trolley is headed straight for them. You are standing some distance off in the train yard, next to a lever. If you pull this lever, the trolley will switch to a different set of tracks. However, you notice that there is one person on the side track. You have two options:
        1. Do nothing and allow the trolley to kill the five people on the main track.
        2. Pull the lever, diverting the trolley onto the side track where it will kill one person.
  4. **What is one of the moral dilemmas presented by the "crash" scenario for driverless cars?**
     1. driverless car with n passengers is proceeding down a road when an obstacle is introduced into its path. If the car hits the obstacle, it will kill the passengers. It can swerve left or right (take evasive maneuvers), which will kill other (m: m >n) bystanders.
        1. Since one design goal of any car is to protect its inhabitants, it seems that the driverless car should swerve and kill the bystanders.
           1. But more people will be killed by the car following its design goal here if it takes evasive maneuvers. How should the car be programmed?
  5. **What are other scenarios (beyond “crash”) that have moral implications?**
     1. 3-car Tailgating with “mixed” Avs & non-AVs
     2. Crosswalk “intervention”
     3. Brake Failure
     4. Chicken between AV & non-AV
  6. **What factors characterize the 6 levels of the SAE J3016 driving automation scheme?**



* 1. **What is the Kantian conception of autonomy? How does it differ from "engineering" autonomy?**
     1. the ability to make morally-relevant choices independent of influencing factors (both nature and nurture). This is how reason becomes “practical.”
  2. **What is Searle's "Chinese Room" argument? How can it apply to the possibility of autonomous (artificial) moral agents?**
     1. Searle imagines himself alone in a room following a computer program for responding to Chinese characters slipped under the door. Searle understands nothing of Chinese, and yet, by following the program for manipulating symbols and numerals just as a computer does, he sends appropriate strings of Chinese characters back out under the door, and this leads those outside to mistakenly suppose there is a Chinese speaker in the room. The narrow conclusion of the argument is that programming a digital computer may make it appear to understand language but could not produce real understanding. Hence the “Turing Test” is inadequate. Searle argues that the thought experiment underscores the fact that computers merely use syntactic rules to manipulate symbol strings, but have no understanding of meaning or semantics.
  3. **What are the conditions for moral agency?**
     1. Moral agency is an individual's ability to make moral judgments based on some notion of right and wrong and to be held accountable for these actions
        1. Decision-making ability
        2. Effective decisions (causal actions)
        3. And perhaps also that the agent intends its actions
  4. **How did we define a social robot?**
     1. A social robot is an autonomous or semi-autonomous robot that interacts and communicates with humans by following the behavioral norms expected by the people with whom the robot is intended to interact.
  5. **What is the basic lesson of the "uncanny valley"?**
     1. 
        1. In [aesthetics](https://en.wikipedia.org/wiki/Aesthetics), the **uncanny valley** is a hypothesized relationship between the degree of an object's resemblance to a human being and the emotional response to such an object. The concept suggests that [humanoid](https://en.wikipedia.org/wiki/Humanoid) objects which imperfectly resemble actual human beings provoke [uncanny](https://en.wikipedia.org/wiki/Uncanny) or strangely familiar feelings of eeriness and revulsion in observers.[[2]](https://en.wikipedia.org/wiki/Uncanny_valley#cite_note-MacDormanIshiguro2006-2) "Valley" denotes a dip in the human observer's affinity for the replica, a relation that otherwise increases with the replica's human likeness.[[3]](https://en.wikipedia.org/wiki/Uncanny_valley#cite_note-MacDormanChattoPadhyay2016-3)
           1. The **uncanny valley** hypothesis predicts that an entity appearing almost human will risk eliciting cold, eerie feelings in viewers.
  6. **What are the characteristics of sociable robot, according to Hanheide?**
     1. embodiment in a situated manner
     2. lifelike qualities (anthropomorphization, tendency to interpret behavior as being intentional)
     3. identifying persons, their actions and intentions: Theory of Mind and empathy core for human awareness
     4. Learning of social situations, shaping robot’s personal history by imitation or mimicry.
     5. being understood: ability to read the activities (expressions, mimic, etc.) of a robot.

***XI. Technology and Moral Responsibility***

* 1. **What are Hart's 4 types of responsibility? Gives examples of each type.**
  2. role
     1. Some professions have special responsibilities in virtue of the role they play in society and the peculiar knowledge that they have.
        1. Attorneys protect lawyer-client confidentiality
        2. Police are authorized to use force to prevent crime
  3. causal
     1. A typical moral agent is able to cause (through rational deliberation) and control his/her actions. A moral patient is someone/thing that matters, morally, when they are the “recipient” of actions.
     2. For Kant, this “causality” is the foundation of morality: autonomy and dignity
  4. capacity
     1. The assumption in the modern era is that only humans have this capacity (but not that all human have it, nor all of the time)
  5. liability

1. Which are “backward” and which are “forward” looking?
   1. What is the role of knowledge in “role” responsibility?
   2. Define culpability, strict liability, and negligence
      1. Culpability: intentional harmful action, w/o justification or excuse, and where likely consequences were known (cf. mens rea)
      2. Strict liability: requirement to pay for damage caused by one’s actions, regardless of fault, negligence, or intent
      3. Negligence: failure to exercise reasonable care
   3. **What types of responsibility are of interest to Kantians, and why?**
      1. Causal and Capacity were of interest to Kant
         1. A typical moral agent is able to cause (through rational deliberation) and control his/her actions. A moral patient is someone/thing that matters, morally, when they are the “recipient” of actions.
         2. For Kant, this “causality” is the foundation of morality: autonomy and dignity
         3. The assumption in the modern era is that only humans have this capacity (but not that all human have it, nor all of the time)
   4. **Be able to explain how responsibility questions arise in every section of this course**
   5. **What is Jonas's imperative of responsibility?**
      1. Modern technology has introduced actions of such novel scale, objects, and consequences that the framework of former ethics can no longer contain them.
      2. He proposes a new categorical imperative:
         1. "Act so that the effects of your action are compatible with the permanence of genuine human life.”
   6. **Explain why Jonas thinks that technology is inherently ambivalent, with respect to ethics**
      1. Technology can be good ‘as such’ but bad when abused, and bad when “good”
      2. B. Technology’s “potential” is employed on ever larger scales as “need”
      3. C. Technology’s magnitude is on temporal and spatial scales
      4. D. Technology forces us to confront anthropocentrism
      5. E. Technology could become apocalyptic
   7. **How has modern technology introduced an "ethical novum" acc. to Jonas?**
      1. The scale and causal range of technological practice…are such that they insert a whole new dimension into the frame of ethical reckoning - a dimension unknown to all former kinds of action. The intrusion of distant future and global scales into our everyday, mundane decisions is an ethical novum which technology has thrust on us; and the ethical category preeminently summoned by this novel fact is: responsibility.”
   8. **What is the import of Jonas's example of plowshares and swords?**
      1. the intentions behind technology don’t matter! The problem of “critical mass”
         1. At first glance it seems easy to distinguish between beneficial and injurious technology by just looking at what the instruments are for.
            1. Plowshares are good, swords are bad: in the messianic age, swords will be beaten into plowshares…But here the vexing dilemma of modern technology leaps to the eye: its “plowshares” can be as disastrous in the long run as its “swords”
   9. **Why does Jonas believe that former ethical systems were anthropocentric?** 
      1. 
   10. **Why acc. to Jonas must we reject anthropocentric ethics in the future?**
       1. But properly understood, the inclusion of the existence of the plenitude as such in the human good, thus the inclusion of its preservation in man's duty, goes beyond the utilitarian and any anthropocentric concern. It allies the human good with the cause of life in general, instead of pitting the one against the other, and grants the latter its own right. Its recognition means that any wanton and needless extinction of species becomes a crime in itself, quite apart from the concurring counsels of intelligent self-interest; and to protect the most nonrenewable and most irreplaceable "re- source" of all, the incredibly rich gene pool deposited by aeons of evolution, becomes a transcendent duty of
   11. **What is Jonas's conception of a "critical mass"?**
       1. Once a "critical mass" is reached in one direction or other, matters may be out of our hands: a positive feedback could take over and start a runaway process in which costs devour benefits in a rising and perhaps irreversible crescendo This is what long-range responsibility must prevent
   12. **What is the main argument of Bloom & Harris concerning AI and our own (human) morality?**
       1. Machines are created to improve the lives of human beings, and one of the attractions of advanced AI is the prospect of robot maids, butlers and chauffeurs (also known as self-driving cars). This is all fine with the sorts of machines we currently have, but as AI improves, we run a moral risk….
       2. For the first time in our history, then, we run the risk of building machines that only monsters would use as they please.”