**List of Concepts to Know for the Midterm**

Intelligence

* What qualifies as intelligent?
  + Natural: Slime mold (solved a maze and figured out shortest path)? Tree? One ant? Ant colony?
  + Artificial: book? Calculator? Bot that plays tic-tac-toe?
* The ability to acquire and apply knowledge and skills. In the context of AI, it refers to the ability of machines or systems to perform tasks that typically require human intelligence.

Emergence

* An interaction of simple elements that create an emergent quality inexplicable at the level of elements.
* Basically can’t see the final product/bigger picture without putting all the little pieces together.
* The phenomenon where complex patterns and behaviors arise from relatively simple interactions among components. In AI, emergence is often observed in systems where intelligence or behavior emerges from the interactions of individual components or agents.

Cognitive revolution

* An intellectual shift in psychology in the 1950s focusing on the internal mental processes driving human behavior (pays attention to language acquisition, memory, problem-solving, and learning)
* A period in the mid-20th century marked by a shift in psychology and related disciplines towards an emphasis on cognitive processes, such as perception, memory, and problem-solving. This revolution influenced the development of AI by providing insights into how human intelligence works.

Symbolic and Connectionist AI

* Symbolic AI: takes as its basis the seemly uniquely human ability to manipulate symbols.
  + Dominant until 1980s
  + Limitations:
    - manually encoding complex systems is hard to manage.
    - GPS effective for very simple games
    - More possibilities = more needed rules
* Connectionist (subsymbolic) AI: takes as its basis the connections of neurons.
  + Became dominant in the early 2010s.
* Two main approaches to AI. Symbolic AI involves manipulating symbols according to rules, while connectionist AI (or neural networks) is based on interconnected nodes that mimic the structure of the human brain.

Expert systems

* Applications of symbolic AI
* Attempted to automate domains of expertise.
  + Knowledge base: consists of all known facts and rules in a domain.
  + Inference engine: system that applies facts and rules to specific case.
* Advantages:
  + Ensures standard performance.
  + inexpensive
* disadvantages:
  + knowledge acquisition very complex task
  + plain weird
* AI systems designed to emulate the decision-making abilities of a human expert in a particular domain by using rules and knowledge representations.

Indian Chessboard

* stupid exponential growth problematic not easy to solve.
* a metaphor used in discussions of exponential growth, often applied to technological progress or the capabilities of AI systems. It refers to the story of a king who was asked to place one grain of rice on the first square of a chessboard, two on the second, four on the third, and so on, doubling the amount each time.

General Problem Solver

* symbolic ai
* solves simple puzzles like tower of Hanoi and missionaries vs cannibals.
* uses code that is human-interpretable.
* A program developed by Herbert Simon and Allen Newell in the late 1950s, designed to simulate human problem-solving abilities. It was an early attempt at creating a system capable of solving a wide range of problems.

Backpropagation

* An algorithm used in training artificial neural networks, which calculates the gradient of the error function with respect to the network’s weights, allowing for the adjustment of those weights to minimize the error.
* Basically shows mistakes in increments of neural networks: catches errors, adjusts mistakes (fixes the weight so error minimizes), and repeats until correct.

The hype cycle

* a model that describes the adoption and maturity of new technologies, characterized by a cycle of hype, disillusionment, and eventual adoption. It is often used to understand the trajectory of emerging technologies, including AI.
* a new thing is presented, people think it’s the most amazing thing in the world (hype part); after a while, it gets boring/ overused/ people think it’s not as amazing as they thought it was (disillusionment), then people start figuring better ways to use this thing/ mature it and they start liking it again (adoption)

AI winters

* periods of reduced funding and interest in AI research and development, typically following periods of overhyped expectations and unmet promises
* the “disillusionment” part of the hype cycle when people get bored of the new product.

Strong v weak AI

* Strong AI = AGI (Artificial General Intelligence)
  + Can meet or exceed human performance on most tasks.
  + Ava from *Ex Machina* is human-like in most ways (robot that looks like a person)
* Weak AI = Narrowly focused on specific tasks
  + Limitations: amazing victories, but limited
    - AlphaGo (game) can’t play chess.
    - Watson (robot) can’t drive a car.
* Path from weak to strong AI?
  + Chatting + chess + driving + identifying.
* Strong AI refers to AI systems that possess general intelligence and can understand, learn, and apply knowledge in a manner distinguishable from humans. Weak AI, also known as narrow AI, refers to systems that are designed for specific tasks and do not possess general intelligence.

Artificial general intelligence (AGI)

* Refers to the hypothetical ability of an artificial intelligence system to understand, learn, and apply knowledge across a wide range of tasks and domains, similar to human intelligence.

The Turing Test

* Sees if someone can differentiate computer vs human.
* A test proposed by Alan Turing to determine a machine’s ability to exhibit intelligent behavior indistinguishable from that of a human. In the test, a human evaluator interacts with both a machine and another human through a text interface and tries to determine which is which.

The Chinese Room

* Chinese person sent messages to John Searle in a room, John had no idea how to read/write/speak Chinese and the girl outside thought he did while he was just using a bunch of graphics to translate and communicate back.
* Shows that computers operate on syntactic manipulation of symbols based on algorithms or programs, but they lack genuine understanding or consciousness like humans.
* A thought experiment proposed by philosopher John Searle to argue against the possibility of strong AI (meet/exceed human performance). It imagines a person inside a room following instructions to manipulate Chinese symbols without understanding the language, similar to how a computer executes programs without understanding their meaning.

AlexNet

* A deep convolutional neural network architecture designed for image classification, which achieved breakthrough performance in the ImageNet Large Scale Visual Recognition Challenge in 2012.

The dominance of deep learning

* The widespread adoption and success of deep learning algorithms, particularly neural networks with many layers, in various AI applications, including computer vision, natural language processing, and speech recognition.

Moore’s Law

* An observation by Gordon Moore that the number of transistors on integrated circuits doubles approximately every two years, leading to exponential growth in computing power and capabilities over time.

**Input, hidden, output layers**

* The layers of nodes in a neural network. The input layer receives input data, the hidden layers process that data through weighted connections and activation functions, and the output layer produces the network’s predictions or outputs.

**Transfer learning**

* A machine learning technique where knowledge gained from solving one problem is applied to a different but related problem, often leading to improved performance and efficiency.

Long tail problem

* A phenomenon observed in distribution curves where a large number of occurrences are clustered at the low end of the distribution (the head), while a smaller number of occurrences extend far out to the right (the tail). In the context of AI, it refers to the challenge of addressing the diverse and less frequent cases or scenarios that are outside the mainstream.

Gradient descent

* An optimization algorithm used in machine learning to minimize the error of a model by adjusting its parameters (weights) iteratively in the direction of steepest descent of the error function.
* Goes hand in hand with backpropagation.

Adversarial AI

* AI systems designed to generate or counteract adversarial inputs, such as images or data, to exploit vulnerabilities or improve robustness.
* “tricks” computer so it can enhance its skills; shows picture of cat but changes a few features so computer thinks it’s a dog.

Explainable AI

* AI systems and algorithms that can provide understandable explanations for their decisions and behavior, particularly in critical or sensitive applications where transparency is important.

Deep fake

* Synthetic media generated by AI algorithms, particularly deep learning techniques, to manipulate or replace audio, video, or images to create realistic but fabricated content.
* Ex: make a celebrity look like they’re doing something they’re not

Natural language processing

* The field of AI focused on enabling computers to understand, interpret and generate human language, including speech recognition, language translation, and sentiment analysis.
* It’s like the computer has a decoder/language translation book to figure out what you’re saying.

IBM Watson

* A cognitive computing system developed by IBM, famous for winning the quiz show Jeopardy in 2011. Watson combines natural language processing, machine learning, and other AI techniques to analyze large volumes of data and answer questions posed in the natural language.

Transformers

* A type of deep learning model architecture based on self-attention mechanisms, originally developed for natural language processing tasks but later applied to various other domains, achieving state-of-the-art results in many cases.
* First used to help computers understand natural language, then used for things like understanding emails, translating languages, and even playing games

Winograd Schemas

* A type of linguistic puzzle designed to test an AI system’s ability to understand and reason about natural language, particularly resolving ambiguous pronouns and references

Data

* Information in a form suitable for processing by a computer. In the context of AI, data is crucial for training, testing, and improving machine learning models and algorithms.

Digital v. analogue data

* Digital data is represented using discrete values, typically binary digits (0s and 1s), while analogue data is represented using continuous values.
  + Digital example: computers – represents data such as texts, images, and numbers, using binary code (0s and 1s). In digital systems, everything is processed step-by-step.
* Analogue systems allow for continuous representation of data
  + Analogue example: vinyl records: the grooves on the record represent the sound wave directly, with changes in depth

The data deluge

* The overwhelming volume, velocity, and variety of data being generated and collected in the modern world, presenting both opportunities and challenges for AI and data-driven decision-making.

Big Data

* Extremely large datasets that are too complex to be processed using traditional data processing applications. Big data often involves high-volume, high-velocity, and high-variety information sources.
* Helps process a lot of information quickly especially for large amounts of data.
* Ex: favorite stores, favorite items at stores, favorite songs

Data science

* An interdisciplinary field that uses scientific methods, algorithms, and systems to extract knowledge and insights from structured and unstructured data.

The singularity

* A hypothetical future event where technological progress, particularly in AI and machine intelligence, accelerates beyond human control or comprehension, leading to unpredictable and potentially profound changes in society.
* Technology will outpower humans at some point.

Superintelligence

* Intelligence that surpasses human intelligence in virtually all domains and activities. Superintelligence is often considered a potential outcome of advanced AI development.

The intelligence explosion

* A concept related to the singularity, suggesting that once AI achieves a certain level of intelligence, it could rapidly self-improve and lead to an exponential increase in intelligence.

Common sense

* The basic level of practical knowledge and reasoning that is generally shared by people within a culture or society. Common sense is often cited as a challenging aspect of AI development, as it involves understanding context, causality, and everyday human experience.

Dimensions of superintelligence

* Various aspects of capabilities that could characterize superintelligent AI, including cognitive abilities, creativity, emotional intelligence, and social understanding.

Paths to superintelligence

* Different theories or approaches to achieving superintelligent AI, including gradual improvement of existing systems, revolutionary breakthroughs, or combinations of multiple approaches.

Solutionism

* The belief that all problems can be solved through technological solutions, often associated with an overly optimistic or simplistic view of technology’s capabilities and implications.

Value alignment problem

* The challenge of ensuring that AI systems and their goals are aligned with human values and preferences, to prevent unintended consequences or conflicts with human interests.