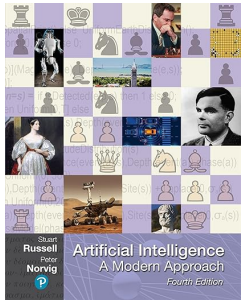
UNIVERSITY OF MARYLAND

MSML610: Advanced Machine Learning

Lesson 01.2: The Foundations of AI

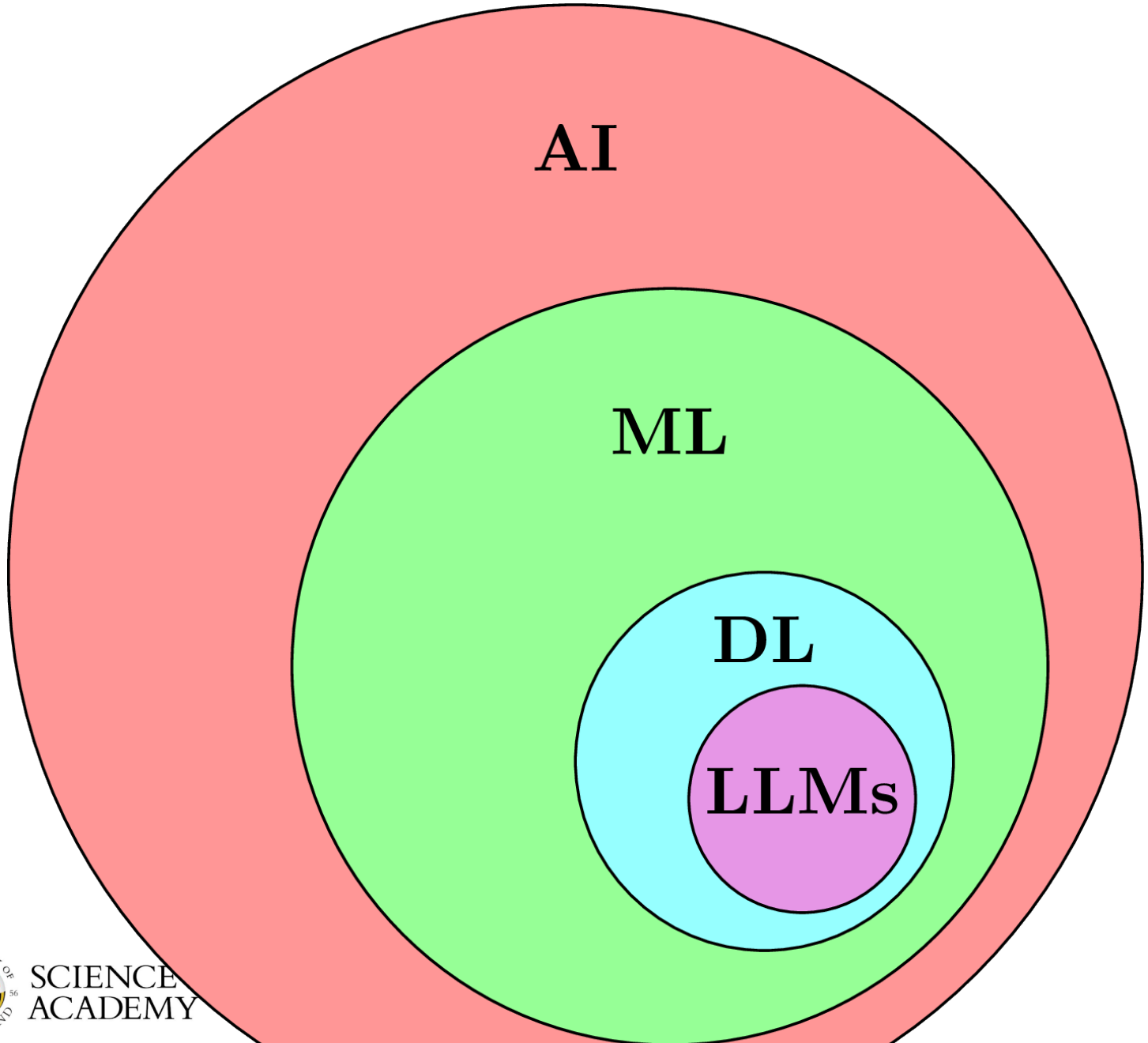
Instructor: Dr. GP Saggese - gsaggese@umd.edu

References: - AIMA (Artificial Intelligence: a Modern Approach), Chap 1



- ***The Foundations of AI***

AI Relates to Many Other Disciplines



SCIENCE ACADEMY

AI and Philosophy (1/2)

- **Can formal rules be used to draw valid conclusions?**
  - **Reasoning**
    - Logic studies rules of proper reasoning
    - Aristotle (400 BCE) formulated laws governing the rational mind
    - Machines were built for arithmetic operations (e.g., Pascaline, 1600)
  - **Rationalism**
    - Use reasoning to understand the world
- **How does the mind arise from a physical brain?**
  - **Dualism**
    - Nature follows physical laws
    - Part of the human mind ("the soul") is exempt from physical laws
  - **Materialism**
    - The mind is a physical system, following the laws of physics
    - Where is free will? Free will is the perception of available choices

AI and Philosophy (2/2)

- **What does knowledge come from?**
  - **Empiricism**
    - Knowledge acquired via senses
    - E.g., learn that trees are green by looking at them
  - **Induction**
    - General rules from associations
    - E.g., many swans are white, infer all swans are white
  - **Logical Positivism**
    - Knowledge as logical theories linked to observations
    - E.g., scientific hypotheses connected to experimental data
- **How does knowledge lead to action?**
  - **Utilitarianism**
    - Actions justified by logic connecting goals and outcomes
  - **Consequentialism**
    - Right or wrong determined by action's expected outcomes
    - E.g., "If you kill, you will go to jail"
  - **Deontological ethics**
    - "Right actions" based on universal laws, not outcomes
    - E.g., "don't kill", "don't lie"

AI and Cognitive Psychology

- **How do humans think and act?**
  - **Cognitive psychology**
    - Brain is an information-processing device
    - Stimuli translated into internal representation
    - Representation manipulated by cognitive processes to derive new internal representations ("beliefs")
    - Representations turned into actions ("goals")
  - **Cognitive science**
    - Use computer models to address memory, language, and logic thinking
    - Dual / opposite of AI
  - **Human-computer interaction**
    - Computers augment human abilities
    - From artificial intelligence (AI) to intelligence augmentation (IA)

AI and Mathematics

- **What are the formal rules to draw valid conclusions?**
  - **Formal logic**
    - Logical deduction rules (Boole, 1850)
    - First-order logic includes objects and relations (Frege, 1879)
  - **Limits to deduction**
    - Some statements are "undecidable"
    - Incompleteness theorem: in any formal theory true statements exist that cannot be proved (Godel, 1931)
- **How do we reason with uncertain information?**
  - **Probability**
    - Mathematics of uncertainty
    - Cardano, Pascal, Bernoulli, Bayes (1500-1700)
  - **Statistics**
    - Combines data with probability
    - E.g., experiment design, data analysis, hypothesis testing, asymptotics

AI and Economics (1/2)

- **How to make decisions to maximize payoff given preferences?**
  - **Economies**
    - Agents maximize economic well-being (utility)
    - Studies desires and preferences
  - **Decision theory**
    - Making decisions under uncertainty for preferred outcomes
    - Probability theory + utility theory
    - E.g., investment choices, policy decisions
- **How to make decisions when payoffs are result of several actions?**
  - **Operations research**
    - Make rational decisions with payoffs for sequence of actions (Bellman, 1957)
    - E.g., Markov Decision Processes
  - **Satisficing**
    - Decisions that are good enough
    - Closer to human behavior
    - E.g., choosing a restaurant that meets basic criteria rather than finding the perfect one

AI and Economics (2/2)

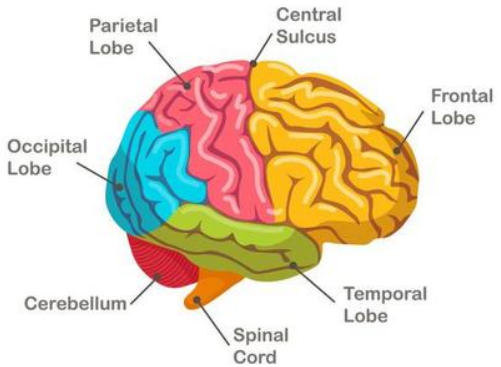
- **How multiple agents with different goals act?**
  - **Large economies**
    - Many agents with no mutual impact
    - Ignore other agents' actions
    - E.g., national economy where individual actions don't affect market
  - **Small economies**
    - One player's actions influence others' utility
    - E.g., local market where one seller's pricing affects competitors
  - **Game theory**
    - Small economies resemble a "game" (Von Neumann, 1944)
    - Rational agents might need randomized strategies
    - E.g., rock-paper-scissors where randomization prevents predictability

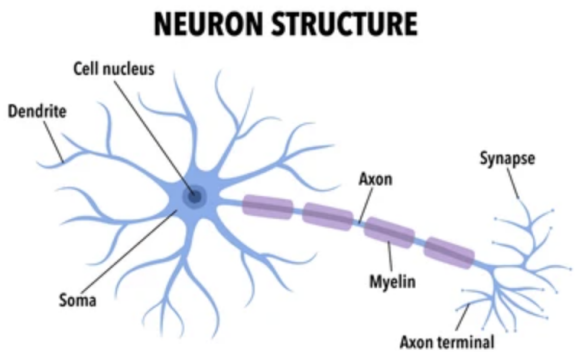
AI and Linguistics

- **How can you create systems that understand natural language?**
  - **Computational linguistics** (NLP)
    - Studies sentence structure and meaning
    - Machine translation (e.g., Google Translate)
    - Sentiment analysis in social media
    - Automated customer support chatbots
- **How does language relate to thought?**
  - **Knowledge representation**
    - How to represent knowledge for computer reasoning
    - E.g., first order knowledge, knowledge graphs

AI and Neuroscience

- **Brain**
  - Parts handle specific cognitive functions
  - Information processing in the cerebral cortex
  - E.g., frontal lobe injury may impair decision-making
- **Anatomy of the brain**
  - Composed of ~100 billion neurons
    - Each neuron connects with 10-100k others via synapses
    - Axons enable long-range connections
  - Signals propagate through electrochemical reactions
  - Short-term pathways support long-term connections (learning)
- **Memory**
  - No theory yet about individual memory storage
  - Current theory: memories reconstructed





The Brain Causes the Mind

- **Simple cells lead to thought and consciousness**
  - **Truly amazing!**
  - Complex processes emerge
- Supercomputers' complexity rivals the brain
- **Brain-machine interface**
  - Brain adjusts to devices
  - E.g., learn to use prosthetics as limbs
- **AI singularity**
  - Future point when AI surpasses human intelligence
    - AI improves autonomously, leading to rapid growth
    - Recursive self-improvement leads to superintelligence
  - Potential societal impact
    - Control problem/value alignment: ensure AI aligns with human values
    - Economic/social disruption due to automation
  - Achieving brain's intelligence level remains unknown

AI and Computer Science

- **What can be computed?**
  - **Algorithm**
    - Procedure to solve problems
    - E.g., algorithm for computing GCD (Euclid, 300 BCE)
  - **Limits to computation**
    - Turing machine (1936): computes any computable function
    - Some functions are non-computable
    - E.g., the halting problem, i.e., decide if a program terminates
  - **Tractability**
    - Complexity classes: polynomial vs exponential complexity
    - Problem is intractable if solving time grows exponentially with size
    - P vs NP

AI and Control Theory

- **How can artifacts operate under their own control?**
  - **Control theory**
    - Study self-regulating feedback control systems
    - E.g., a water regulator that maintains a constant water flow
    - Mechanisms to minimize error between current and goal states
    - Kalman Filter (Kalman, 1960)
    - Based on calculus, matrix, stochastic optimal control
    - AI: logical inference, symbolic planning, computation

AI and Computer Engineering

- **How can we build an efficient compute**
  - **Electronic computers**
    - Built during World War II
  - **Moore's Law**
    - Performance doubled every 18 months (1970-2005)
    - Power and scaling issues shifted focus to multi-core over clock speed
  - **Hardware for AI**
    - GPUs
    - TPUs
    - Wafer-scale engines
  - **Current Trends**
    - Massive parallelism (like brain function)
    - Computing power doubling every 3 months
    - GPUs / TPUs used in deep learning
    - High precision (e.g., 64b) often unnecessary
  - **Quantum Computing**