Artificial General Intelligence

1. Objective and Strategy

Dr. Pei Wang

Dr. Patrick Hammer

Dr. Robert Johansson

Digital Futures (KTH, SU, RISE)

Stockholm, Fall 2022

Artificial Intelligence (AI)

- Computers can do mental work
- AI: to build mind-like computers
- It turned out to be very hard, so the mainstream turned to specific problems and functions
- The same is the case for machine learning, which approximates or optimizes specific functions
- Intelligence may not be a set of special-purpose functions or skills

Artificial General Intelligence (AGI)

- AGI: To take AI as one problem, like the original "AI", and similar to "human-level AI", "strong AI"
- AGI started about 20 years ago as a small research community, with its conference and journal, though there is still no common theory, technique, even goal
- This course will use one concrete AGI project as a sample to explore the problems in the field

Specifying the objective

- "Intelligence" is an abstraction of human intelligence, and a graded category
- Perspectives: structure, behavior, capabilities, functions, principle
- Extreme positions: anthropocentrism and panpsychism
- "AI Effect": What is the relation between "intelligence" and "computation"?

Forms of intelligence

Widely acknowledged forms of intelligence:

- Human intelligence
- Artificial (computer) intelligence
- Animal intelligence
- Group (collective) intelligence
- Alien (extraterrestrial) intelligence

What is shared by them?

My working definition

Intelligence is the ability of adapting to the environment while operating with insufficient knowledge and resources

- The Assumption of Insufficient Knowledge and Resources (AIKR): *finite*, *real-time*, *open*
- Adapting: using the *past* for the *current* and the *future*; using the *bounded supplies* for the *unbounded demands*

Contrasting with brain modeling

- Rationale: The human brain is the best-known mechanism that creates intelligence
- Example: Thousand-Brains Theory
- Challenges: difficulty, necessity, generality
- To model the brain or the mind?
- Flying vs. flying like a bird
- Brain as source of inspiration vs. justification

Contrasting with behaviorism

- Rationale: Intelligent systems are those that are indistinguishable from human in verbal behaviors
- Example: Turing Test
- Challenges: difficulty, necessity, generality
- Why should an AI pretend to be a human?
- Adaptive: behaviors depend on experience, while AI cannot have human experience
- To selectively inherit from psychology

Contrasting with problem solving

- Rationale: Intelligent systems are those that can solve hard problems
- Example: Expert systems
- Challenges: identity, originality, flexibility
- Which problems? Whose intelligence?
- Adaptive means variable problem-solving processes, rather than a specific set of skills
- Theoretical model with application potentials

Contrasting with function collection

- Rationale: Intelligence includes reasoning, learning, planning, problem-solving, etc.
- Example: AI textbooks
- Challenges: commonality, coherency
- "AI completeness"
- Phenomena vs. mechanism
- Cognitive functions as aspects of the same adaptation process

Contrasting with perfect rationality

- Rationale: Intelligence is the ability to do the right thing
- Example: Classical models of rationality
- Challenges: over-idealization, applicability
- As a realistic restriction, AIKR changes the problems fundamentally
- Relative rationality: to do the right thing with respect to the available knowledge and resources

Strategies of building AGI

- Hybrid: linking multiple existing AI techniques into one system
- Integrated: creating an architecture with multiple functional modules
- Unified: extending a single technique to cover various functions

Representative AGI projects

goal\path	hybrid	integrated	unified
principle			AERA, AIXI, NARS
function		OpenCog, Sigma, Soar	SNePS
capability			Cyc, Gato
behavior		ACT-R, LIDA, MicroPsi	
structure			HTM

AGI projects classified by objective and strategy

Suggested Readings

- Alan Turing, Computing machinery and intelligence
- Yann LeCun, A Path Towards Autonomous Machine Intelligence
- Pei Wang, Artificial General Intelligence — A gentle introduction
- Pei Wang, Non-Axiomatic Logic, Chapter 1
- Pei Wang, Rigid Flexibility, Chapter 1,2