

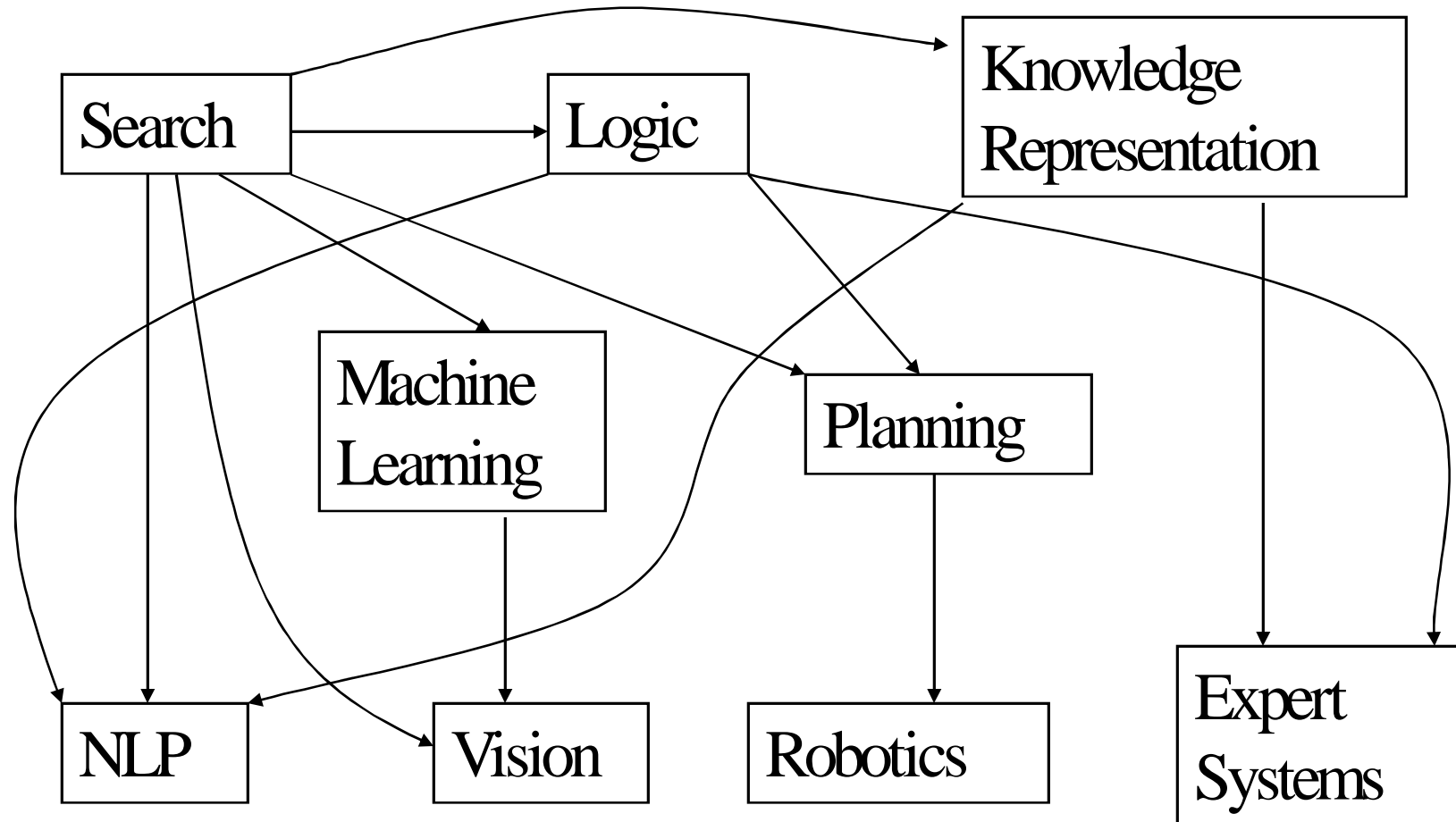
# CS8691-ARTIFICIAL INTELLIGENCE

# Course Learning Outcomes

At the end of this course:

- Knowledge and understanding  
You should have a knowledge and understanding of the basic concepts of Artificial Intelligence including Search, Game Playing, KBS (including Uncertainty), Planning and Machine Learning.
- Intellectual skills  
You should be able to use this knowledge and understanding of appropriate principles and guidelines to synthesise solutions to tasks in AI and to critically evaluate alternatives.
- Practical skills  
You should be able to use a well known declarative language (Prolog) and to construct simple AI systems.
- Transferable Skills  
You should be able to solve problems and evaluate outcomes and alternatives

# Areas of AI and Some Dependencies



# What is Artificial Intelligence ?

- making computers that think?
- the automation of activities we associate with human thinking, like decision making, learning ... ?
- the art of creating machines that perform functions that require intelligence when performed by people ?
- the study of mental faculties through the use of computational models ?

# What is Artificial Intelligence ?

- the study of computations that make it possible to perceive, reason and act ?
- a field of study that seeks to explain and emulate intelligent behaviour in terms of computational processes ?
- a branch of computer science that is concerned with the automation of intelligent behaviour ?
- anything in Computing Science that we don't yet know how to do properly ? (!)

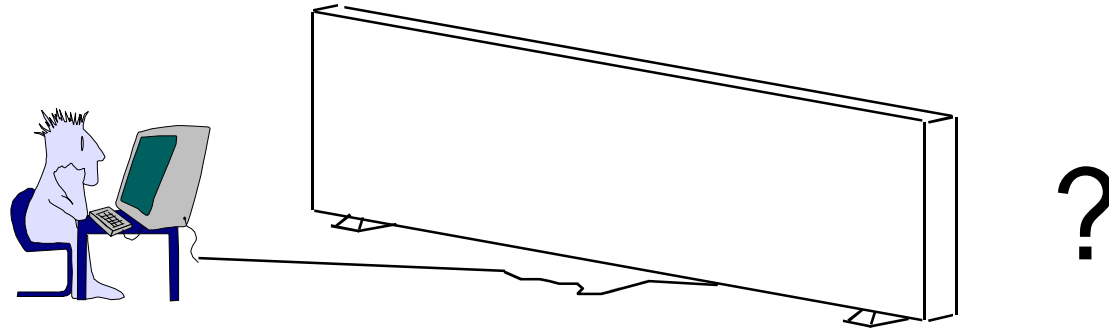
# What is Artificial Intelligence ?

<b>THOUGHT</b>	<b>Systems that think like humans</b>	<b>Systems that think rationally</b>
<b>BEHAVIOUR</b>	<b>Systems that act like humans</b>	<b>Systems that act rationally</b>
	<b>HUMAN</b>	<b>RATIONAL</b>

# Systems that act like humans: Turing Test

- “The art of creating machines that perform functions that require intelligence when performed by people.” (Kurzweil)
- “The study of how to make computers do things at which, at the moment, people are better.” (Rich and Knight)

# Systems that act like humans



- You enter a room which has a computer terminal. You have a fixed period of time to type what you want into the terminal, and study the replies. At the other end of the line is either a human being or a computer system.
- If it is a computer system, and at the end of the period you cannot reliably determine whether it is a system or a human, then the system is deemed to be intelligent.



# Systems that act like humans

- The Turing Test approach
  - a human questioner cannot tell if
    - there is a computer or a human answering his question, via teletype (remote communication)
  - The computer must behave intelligently
- Intelligent behavior
  - to achieve human-level performance in all cognitive tasks

# Systems that act like humans

- These cognitive tasks include:
  - *Natural language processing*
    - for communication with human
  - *Knowledge representation*
    - to store information effectively & efficiently
  - *Automated reasoning*
    - to retrieve & answer questions using the stored information
  - *Machine learning*
    - to adapt to new circumstances

# The total Turing Test

- Includes two more issues:
  - *Computer vision*
    - to perceive objects (seeing)
  - *Robotics*
    - to move objects (acting)

# What is Artificial Intelligence ?

<b>THOUGHT</b>	<b>Systems that think like humans</b>	<b>Systems that think rationally</b>
<b>BEHAVIOUR</b>	<b>Systems that act like humans</b>	<b>Systems that act rationally</b>
	<b>HUMAN</b>	<b>RATIONAL</b>

# Systems that think like humans: cognitive modeling

- Humans as observed from 'inside'
- How do we know how humans think?
  - Introspection vs. psychological experiments
- Cognitive Science
- "The exciting new effort to make computers think . . . machines with *minds* in the full and literal sense" (Haugeland)
- "[The automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning . . ." (Bellman)

# What is Artificial Intelligence ?

**THOUGHT**

**Systems that think  
like humans**

**Systems that think  
rationally**

**BEHAVIOUR**

**Systems that act  
like humans**

**Systems that act  
rationally**

**HUMAN**

**RATIONAL**

# Systems that think 'rationally'

## "laws of thought"

- Humans are not always 'rational'
- Rational - defined in terms of logic?
- Logic can't express everything (e.g. uncertainty)
- Logical approach is often not feasible in terms of computation time (needs 'guidance')
- "The study of mental facilities through the use of computational models" (Charniak and McDermott)
- "The study of the computations that make it possible to perceive, reason, and act" (Winston)

# What is Artificial Intelligence ?

<b>THOUGHT</b>	<b>Systems that think like humans</b>	<b>Systems that think rationally</b>
<b>BEHAVIOUR</b>	<b>Systems that act like humans</b>	<b>Systems that act rationally</b>
	<b>HUMAN</b>	<b>RATIONAL</b>



# Systems that act rationally: “Rational agent”

- **Rational** behavior: doing the right thing
- **The right thing**: that which is expected to maximize goal achievement, given the available information
- Giving answers to questions is ‘acting’.
- I don't care whether a system:
  - replicates human thought processes
  - makes the same decisions as humans
  - uses purely logical reasoning

# Systems that act rationally

- Logic → only *part* of a rational agent, not *all* of rationality
  - Sometimes logic cannot reason a correct conclusion
  - At that time, some *specific (in domain) human knowledge* or information is used
- Thus, it covers more generally different situations of problems
  - Compensate the incorrectly reasoned conclusion

# Systems that act rationally

- Study AI as rational agent –

2 advantages:

- It is more general than using logic only
  - Because: LOGIC + Domain knowledge
- It allows extension of the approach with more scientific methodologies

# Rational agents

- An **agent** is an entity that perceives and acts
- This course is about designing rational agents
- Abstractly, an agent is a function from percept histories to actions:

$$[f: P^* \rightarrow A]$$

- For any given class of environments and tasks, we seek the agent (or class of agents) with the best performance
- Caveat: computational limitations make perfect rationality unachievable
  - → design best program for given machine resources

- **Artificial**
  - Produced by human art or effort, rather than originating naturally.
- **Intelligence**
- is the ability to acquire knowledge and use it" [Pigford and Baur]
- **So AI was defined as:**
  - **AI** is the study of ideas that enable computers to be intelligent.
  - **AI** is the part of computer science concerned with design of computer systems that exhibit human intelligence(From the Concise Oxford Dictionary)

From the above two definitions, we can see that AI has two major roles:

- Study the intelligent part concerned with humans.
- Represent those actions using computers.