

# Lecture 1: Introduction to Artificial Intelligence; Intelligent Agents

ICS 171, Summer 2000

ICS-171:Lecture 1: 1

## Course Details

- **Instructor**
  - Stephen Bay
  - sbay@ics.uci.edu
  - office CS/E 305
  - office hours 1-2pm TuTh (or by appointment)
- **TAs**
  - Suraj Nagasrinivasa, surajn@ics.uci.edu
  - office?
  - office hours?
- **Text and Notes**
  - *Artificial Intelligence: A Modern Approach*, Stuart Russell and Peter Norvig, Prentice Hall, 1995.
  - notes and handouts on web page
- **Course Web Page:** <http://www.ics.uci.edu/~sbay/ics171.html>
- **Newsgroup:** ics.171
- **Lab:** CS 364 8am-5pm

ICS-171:Lecture 1: 2

## Grading

- **Grading Basis**
  - 15% homework + 15% midterm + 20% project + 50% final
  - 20% project + 80% final
- **Late Policy**
  - no late homeworks, -20%/day for project
- **Graded Quizzes**
  - can be picked up in the Distribution Center or in Discussion
- **Regrading**
  - turn in paper within 1 week of receiving it with a written note explaining why you are requesting a regrade
  - Note that regrading may result in your mark *increasing or decreasing*
- **Bonus Points**
  - I will occasionally offer bonus questions to students. I may announce other opportunities for bonus points during the course.

ICS-171:Lecture 1: 3

## Academic Honesty

- **It is your responsibility to be familiar with UCI's policies on academic honesty.**
  - <http://www.editor.uci.edu/catalogue/appx/appx.2.html>
  - [http://www.ics.uci.edu/~ucounsel/continuing\\_students/cheat.html](http://www.ics.uci.edu/~ucounsel/continuing_students/cheat.html)
- **Violations can result in an F and a letter in your permanent student file**
- **All homework assignments and the project must be done individually**
  - you are encouraged to discuss the problem at a high level with others, but your solution must be your own

ICS-171:Lecture 1: 4

## Academic Honesty

### Cheating (from ICS Policy Page)

- Turning in someone else's work as your own (with or without his or her knowledge). Turning in a completely duplicated assignment is a flagrant offense, but even copying only a portion of the assignment and turning it in as your own is considered cheating.
- Allowing someone else to turn in your work as his or her own.
- Several people writing one program and turning in multiple copies, all represented (implicitly or explicitly) as individual work.
- Using any part of someone else's work without the proper acknowledgment.
- Stealing an examination or a solution from the instructor. This is an extremely flagrant offense

ICS-171:Lecture 1: 5

## Outline of Syllabus

- **Introduction to AI**
- **Intelligent Agents**
- **Search**
  - Problem solving as search. Uninformed and informed search methods. Iterative improvement algorithms. Game playing.
- **Reasoning**
  - Knowledge representation and inference. Propositional logic, predicate logic, reasoning under uncertainty, rule-based systems.
- **Learning**
  - Improving the computer with experience. Concept learning from examples: decision trees, rule learners, naïve Bayesian classifiers, perceptrons and neural networks.
- **Representation and Abstraction**

ICS-171:Lecture 1: 6

## What is AI?

ICS-171:Lecture 1: 7

## Today's Lecture Outline

- **What is Artificial Intelligence?**
  - AI is the technology of making computers more intelligent
  - so what is “intelligence”?
- **A look at a fictional AI system:**
  - The “character” HAL in the movie 2001 (made in 1969)
  - what predictions were made for HAL?
  - how have these turned out?
  - what makes AI problems hard?
- **Real-World Applications of AI**
  - AI is alive and well in various “every day” applications
    - many products, systems, have AI components
- **Reading assignment:** Chapter 1 & 2 in R&N,
  - for philosophical background, see also Chapter 26

ICS-171:Lecture 1: 8

## What is Intelligence?

- **Intelligence:**
  - “the capacity to learn and solve problems” (Webster’s dictionary)
  - in particular, *the ability*
    - *to solve novel problems*
    - *to act rationally*
    - *to act like humans*
    - *to acquire knowledge, learn from experience*
  - Modern AI mainly models how **ideal** agents should act
- **Artificial Intelligence**
  - build and understand intelligent entities or agents
  - 2 main approaches: “engineering” versus “cognitive modeling”
  - involves ideas from:
    - philosophy, psychology, cognitive science, computer science
    - mathematics, engineering

ICS-171:Lecture 1: 9

## What’s involved in Intelligence?

- **Ability to interact with the real world**
  - to perceive, understand, and act
  - e.g., speech recognition and understanding and synthesis
  - e.g., image understanding (e.g. fingerprints)
  - e.g., ability to take actions, have an effect
- **Reasoning and Planning**
  - modeling the external world, given input
  - solving new problems, planning, and making decisions
  - ability to deal with unexpected problems, uncertainties
- **Learning and Adaptation**
  - we are continuously learning and adapting
  - our internal models are always being “updated”
    - e.g., a baby learning to categorize and recognize animals

ICS-171:Lecture 1: 10

## HAL: from the movie 2001

- **2001: A Space Odyssey**
  - classic science fiction movie from 1969
- **HAL**
  - part of the story centers around an intelligent computer called HAL
  - HAL is the “brains” of an intelligent spaceship
  - in the movie, HAL can
    - speak easily with the crew
    - see and understand the emotions of the crew
    - navigate the ship automatically
    - diagnose on-board problems
    - make life-and-death decisions
    - display emotions
- **In 1969 this was science fiction: is it still science fiction?**

ICS-171:Lecture 1: 11

## Could we build HAL's hardware in 2001?

- **How complicated is our brain?**
  - a neuron, or nerve cell, is the basic information processing unit
  - estimated to be on the order of  $10^{12}$  neurons in a human brain
  - many more synapses ( $10^{14}$ ) connecting these neurons
  - cycle time:  $10^{-3}$  seconds (1 millisecond)
- **How complex can we make computers?**
  - $10^6$  or more transistors per CPU
  - supercomputer: hundreds of CPUs,  $10^9$  bits of RAM
  - cycle times: order of  $10^{-8}$  seconds
- **Conclusion**
  - **YES:** in the near future we can have computers with as many basic processing elements as our brain, but with
    - far fewer interconnections (wires or synapses) than the brain
    - much faster updates than the brain
  - **but** building hardware is very different from making a computer behave like a brain!

ICS-171:Lecture 1: 12

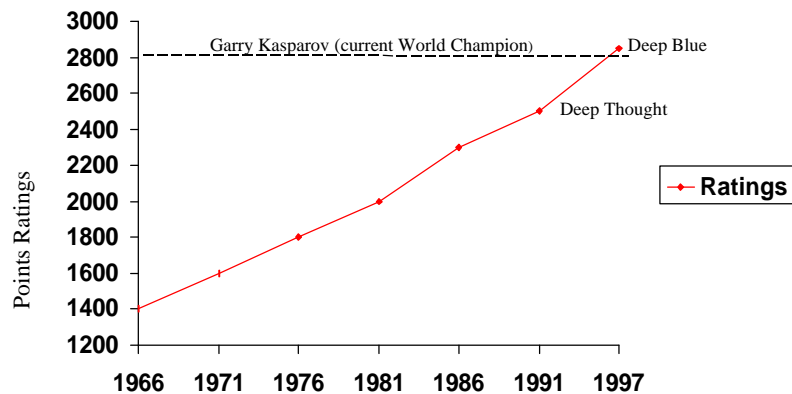
## Could HAL be foolproof in 2001?

- A “foolproof” system is one that never makes an error:
  - Types of possible computer errors
    - hardware errors, e.g., memory errors
    - software errors, e.g., coding bugs
    - “human-like” errors
  - Clearly, hardware and software errors are possible in practice
  - what about “human-like” errors?
- An intelligent system can make errors and still be intelligent
  - humans are not right all of the time
  - we learn and adapt from making mistakes
    - e.g., consider learning to surf or ski
      - we improve by taking risks and falling
      - an intelligent system can learn in the same way
- Conclusion:
  - **NO**: intelligent systems will not (and need not) be foolproof

ICS-171:Lecture 1: 13

## Could HAL beat the crew at chess in 2001 ?

- Chess Playing is a classic AI problem
  - well-defined problem
  - very complex: difficult for humans to play well



- Conclusion: **YES**: today’s computers can beat even the best human

ICS-171:Lecture 1: 14

## Could HAL talk and sound like a human in 2001?

- **This is known as “speech synthesis”**
  - translate text to phonetic form
    - e.g., “fictitious” -> fik-tish-es
  - use pronunciation rules to map phonemes to actual sound
    - e.g., “tish” -> sequence of basic audio sounds
- **Difficulties**
  - sounds made by this “lookup” approach sound unnatural
  - sounds are not independent
    - e.g., “act” and “action”
    - modern systems (e.g., at AT&T) can handle this pretty well
  - a harder problem is emphasis, emotion, etc
    - humans understand what they are saying
    - machines don't: so they sound unnatural
- **Conclusion: NO**, for complete sentences, but YES for individual words

ICS-171:Lecture 1: 15

## Could HAL recognize human speech in 2001?

- **Speech Recognition:**
  - mapping sounds from a microphone into a list of words
  - classic problem in AI, very difficult
    - “Lets talk about how to wreck a nice beach”
    - (I really said “Lets talk about how to recognize speech”)
- **Recognizing single words from a small vocabulary**
  - systems can do this with high accuracy (order of 99%)
  - e.g., directory inquiries
    - limited vocabulary (area codes, city names)
    - computer tries to recognize you first, if unsuccessful hands you over to a human operator
    - saves millions of dollars a year for the phone companies

ICS-171:Lecture 1: 16

## Could HAL recognize human speech in 2001? (ctd.)

- **Recognizing normal speech is much more difficult**
  - speech is continuous: where are the boundaries between words?
  - large vocabularies
    - can be many thousands of possible words
    - we can use **context** to help figure out what someone said
      - e.g., hypothesize and test
      - try telling a waiter in a restaurant:  
“I would like some cream and sugar in my coffee”
  - background noise, other speakers, accents, clogs, etc
  - on normal speech, modern systems are only about 60% accurate
- **Conclusion: NO**, normal speech is too complex to accurately recognize, but **YES** for restricted problems
  - (e.g., recent software for PC use by IBM, Dragon systems, etc)

ICS-171:Lecture 1: 17

## Could HAL understand what we are saying in 2001?

- **Understanding is different to recognition:**
  - “Time flies like an arrow”
    - assume the computer can recognize all the words
    - but how could it understand it?
      - 1. time passes quickly like an arrow?
      - 2. command: time the flies the way an arrow times the flies
      - 3. command: only time those flies which are like an arrow
      - 4. “time-flies” are fond of arrows
    - only 1. makes any sense, but how could a computer figure this out?
      - clearly humans use a lot of implicit commonsense knowledge in communication
- **Conclusion: NO**, much of what we say is beyond the capabilities of a computer to understand at present

ICS-171:Lecture 1: 18

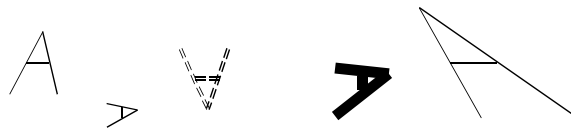
## Could HAL learn to adapt in 2001?

- **Learning and Adaptation**
  - consider a computer learning to drive on the freeway
  - we could teach it lots of rules about what to do
  - or we could let it drive and steer it back on course when it heads for the embankment
    - e.g., RALPH at CMU
      - recently drove 98% of the way from Pittsburgh to San Diego without any human assistance
  - **machine learning** allows computers to learn to do things without explicit programming
  - many successful applications:
    - requires some “set-up”: does not mean your PC can learn to forecast the stock market or become a brain surgeon
- **Conclusion: YES**, computers can learn and adapt, when presented with information in the appropriate way

ICS-171:Lecture 1: 19

## Could HAL see in 2001?

- **Recognition v. Understanding (like Speech)**
  - Recognition and Understanding of Objects in a scene
    - look around this room
    - you can effortlessly recognize objects
    - human brain can map 2d visual image to 3d “map”
- **Why is visual recognition a hard problem?**



- **Conclusion: mostly NO**: computers can only “see” certain types of objects under limited circumstances: but **YES** for certain constrained problems (e.g., face recognition)

ICS-171:Lecture 1: 20

## Could HAL plan and make decisions in 2001?

- **Intelligence**
  - involves solving problems and making decisions and plans
  - e.g., you want to visit your cousin in Boston
    - you need to decide on dates, flights
    - you need to get to the airport, etc
    - involves a sequence of decisions, plans, and actions
- **What makes planning hard?**
  - the world is not predictable:
    - your flight is canceled or there's a backup on the 405
  - there are a potentially huge number of details
    - do you consider all flights? all dates?
      - no: commonsense constrains your solutions
  - AI systems are only successful in constrained planning problems
- **Conclusion: NO**, real-world planning and decision-making is still beyond the capabilities of modern computers
  - exception: very well-defined, constrained problems

ICS-171:Lecture 1: 21

## Summary of State of AI Systems in Practice

- **Speech synthesis, recognition and understanding**
  - very useful for limited vocabulary applications
  - unconstrained speech understanding is still too hard
- **Computer vision**
  - works for constrained problems (hand-written zip-codes)
  - understanding real-world, natural scenes is still too hard
- **Learning**
  - adaptive systems are used in many applications: have their limits
- **Planning and Reasoning**
  - only works for constrained problems: e.g., chess
  - real-world is too complex for general systems
- **Overall:**
  - many components of intelligent systems are “doable”
  - there are many interesting research problems remaining

ICS-171:Lecture 1: 22

## Website of the Day

- <http://mitpress.mit.edu/e-books/Hal/>
- **HAL's Legacy: 2001's Computer as Dream and Reality**
  - MIT Press, 1997, David Stork (ed.)
  - discusses
    - HAL as an intelligent computer
    - are the predictions for HAL realizable with AI today?
- **The website contains**
  - full text and abstracts of chapters from the book
  - links to related material and AI information
  - sound and images from the film

ICS-171:Lecture 1: 23

## Intelligent Systems in Your Everyday Life

- **Post Office**
  - automatic address recognition and sorting of mail
- **Banks**
  - automatic check readers, signature verification systems
  - automated loan application classification
- **Telephone Companies**
  - automatic voice recognition for directory inquiries
  - automatic fraud detection,
  - classification of phone numbers into groups
- **Credit Card Companies**
  - automated fraud detection, automated screening of applications
- **Computer Companies**
  - automated diagnosis for help-desk applications

ICS-171:Lecture 1: 24

## AI Applications: Consumer Marketing

- Have you ever used any kind of credit/ATM/store card while shopping?
  - if so, you have very likely been “input” to an AI algorithm
- All of this information is recorded digitally
- Companies like Nielsen gather this information weekly and search for patterns
  - general changes in consumer behavior
  - tracking responses to new products
  - identifying customer segments: targeted marketing, e.g., they find out that consumers with sports cars who buy textbooks respond well to offers of new credit cards.
  - Currently a very hot area in marketing
- **How do they do this?**
  - Algorithms (“data mining”) search data for patterns
  - based on mathematical theories of learning
  - completely impractical to do manually

ICS-171:Lecture 1: 25

## AI Applications: Identification Technologies

- **ID cards**
  - e.g., ATM cards
  - can be a nuisance and security risk:
    - cards can be lost, stolen, passwords forgotten, etc
- **Biometric Identification**
  - walk up to a locked door
    - camera
    - fingerprint device
    - microphone
  - computer uses your biometric signature for identification
    - face, eyes, fingerprints, voice pattern
- **How does this work?**
  - compare data from person at door with stored library
  - how does one define similarity? (say you have a cold, or got a haircut?)
  - learning algorithms can learn the matching process by analyzing a large library database off-line, can improve its performance.

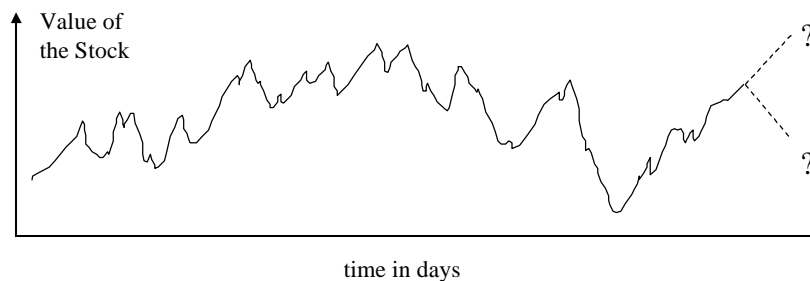
ICS-171:Lecture 1: 26

## AI Applications: Intrusion Detection

- **Computer security**
  - we each have specific patterns of computer use
    - times of day
    - lengths of sessions
    - command used, sequence of commands, etc
  - would like to learn the “signature” of each authorized user
  - can identify non-authorized users
- **How can the program automatically identify users?**
  - record user’s commands and time intervals
  - characterize the patterns for each user
  - model the variability in these patterns
  - classify (online) any new user by similarity to stored patterns
- same ideas being used by Microsoft (and other companies) for “user modeling”

ICS-171:Lecture 1: 27

## AI Applications: Predicting the Stock Market



- **The Prediction Problem**
  - given the past, predict the future
  - very difficult problem!
  - we can use learning algorithms to learn a predictive model from historical data
    - $\text{prob}(\text{increase at day } t+1 \mid \text{values at day } t, t-1, t-2, \dots, t-k)$
  - such models are routinely used by banks and financial traders to manage portfolios worth millions of dollars

ICS-171:Lecture 1: 28

## AI-Applications: Machine Translation

- **Language problems in international business**
  - e.g., at a meeting of Japanese, Korean, Vietnamese and Swedish investors, no common language
  - or: you are shipping your software manuals to 127 countries
  - solution; hire translators to translate
  - would be much cheaper if a machine could do this!
- **How hard is automated translation**
  - very difficult!
  - e.g., English to Russian
    - “The spirit is willing but the flesh is weak” (English)
    - “the vodka is good but the meat is rotten” (Russian)
  - not only must the words be translated, but their meaning also!
- **Nonetheless....**
  - commercial systems can do a lot of the work very well (e.g., restricted vocabularies in software documentation)
  - algorithms which combine dictionaries, grammar models, etc.
  - see for example [babelfish.altavista.com](http://babelfish.altavista.com)

ICS-171:Lecture 1: 29

## Summary

- **Artificial Intelligence involves the study of:**
  - automated recognition and understanding of speech, images, etc
  - learning and adaptation
  - planning, reasoning, and decision-making
- **AI has made substantial progress in**
  - recognition and learning
  - some planning and reasoning problems
- **AI Applications**
  - improvements in hardware and algorithms => AI applications in industry, finance, medicine, and science.
- **AI Research**
  - many problems still unsolved: AI is a fun research area!

ICS-171:Lecture 1: 30

## Intelligent Agents

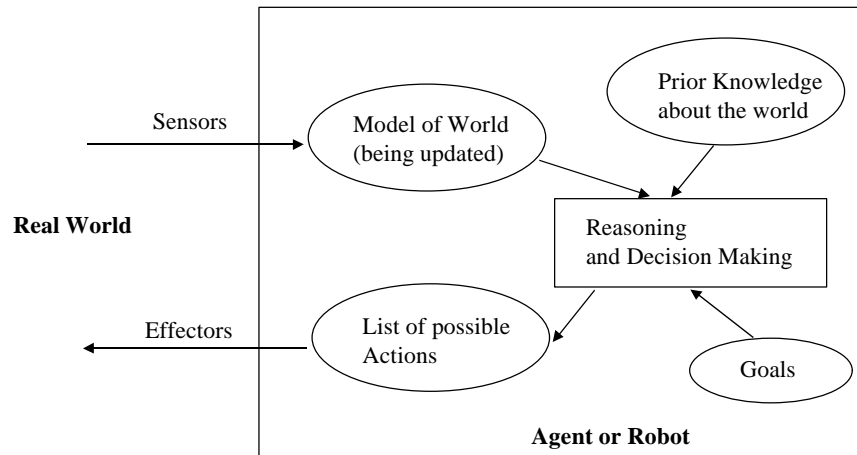
ICS-171:Lecture 1: 31

## What is an Agent?

- An agent perceives its environment with sensors and then acts upon the environment with its effectors to achieve its goals
- Agent metaphor provides a useful framework for thinking about and designing artificial intelligence systems
  - percepts (inputs)
  - actions (outputs)
  - goals
  - environment

ICS-171:Lecture 1: 32

## A Simple Intelligent Agent Architecture



ICS-171:Lecture 1: 33

## Examples of Intelligent Agents

Domain	Inputs	Output/Action	Goal
Chess-Playing	Current board position	Next move	Win game
Web-search	Keywords, Web-pages	List of Web pages	Find relevant Web pages
Speech	Audio signal (microphone)	List of words	Recognition
Driving	Images, signals, position	Movement	Drive from A to B
and many more.....			

ICS-171:Lecture 1: 34

## Agent Concepts (1)

- **Basic cycle:**
  - perceive information => reason/decide => act, repeat.....
- **How should an agent act? A Rational Agent**
  - Performance/utility/success function, is a quantitative measure of how successful an agent is
    - it measures how successful the agent is
  - A rational agent always tries to maximize its success function
  - examples of success functions
    - final grade in ICS 171: to maximize it, agent decides to study
    - number of faces in an image recognized correctly
    - length of time taken to navigate across campus (robot)
    - total time of idle aircraft (minimize it): airline planning

ICS-171:Lecture 1: 35

## Agent Concepts (2)

- **In practice there is often uncertainty**
  - e.g., agent studies hard => expected success is high, but nonetheless final exam questions are unusually hard
  - e.g. agent forecasts the stock market: forecasts tomorrow's direction correctly 60% of the time but incorrect 40% of time
  - => make decision which maximizes success on average
    - (we will return to this later in the quarter)
- **Simplifications we will assume for now:**
  - environment is static:
    - i.e., information does not change with time
  - there is no uncertainty
    - the agent knows everything it needs to know about the state of the world

ICS-171:Lecture 1: 36

## Agent Concepts (3)

- **Types of Agent Architectures**
  - lookup table:
    - matches all perception sequences to actions: impossibly large!!
  - generalize to reflex/reactive agents:
    - summarizes table as rules: *if input A and B, then do action C*
    - problem: lacks information about the state of the world
  - add models of the external world
    - try to track the state of the world
    - e.g., when driving agent assumes physics of how objects move without explicitly looking at each instant
  - add goal-driven behavior:
    - e.g., at an intersection which direction to go?
    - work backwards from goal: less efficient but more flexible
- **We will look at specific goal-driven problems (i.e., state-space search)**

ICS-171:Lecture 1: 37

## Summary

- An **agent** perceives and acts in an environment to achieve specific goals
- **Agents** can be characterized by their
  - percepts
  - actions
  - goals
  - environment
- **Agent Types:**
  - look up table
  - reflex
  - goal based
  - utility based

ICS-171:Lecture 1: 38