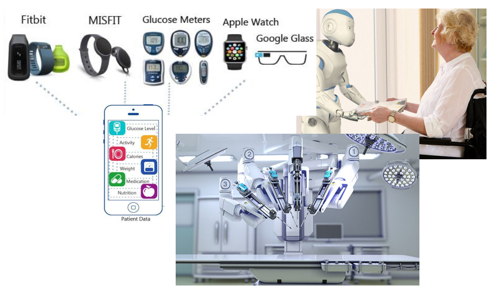
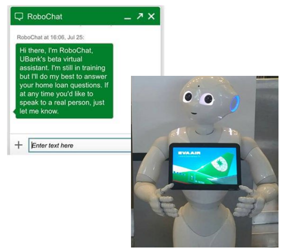
COMP30024 Artificial Intelligence

AI is Everywhere

Healthcare

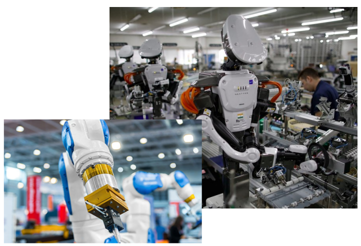


Customer Service

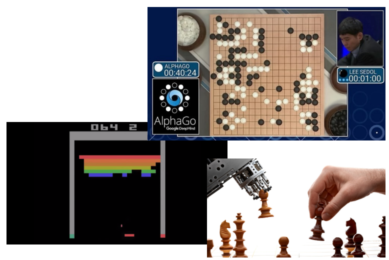


Transportation

Manufacturing

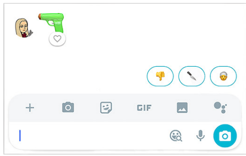


Gaming

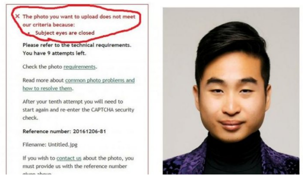


Smart Homes

Examples of AI Failure

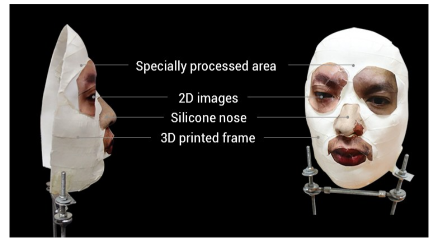


Google Allo responds to a gun emoji with a turban emoji



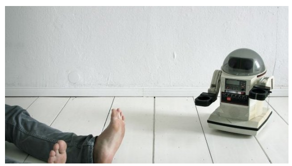
Robot passport checker rejects Asian man’s application because “eyes are closed”

Amazon Alexa starts a party, and the neighbours call the cops



Hackers broke FaceID after a week iphone X release

Autonomous van in accident on its first day



Kid-friendly robot goes crazy and injures a young boy

Street sign hack fools self-driving cars

Classifying all these images as mushroom with 100% confidence

Want to understand AI bounds? Learn from its failures.

Our AI Team

◊ Lecturers: Dr. Sarah Erfani (`sarah.erfani@unimelb.edu.au`)
Prof. Chris Leckie (`caleckie@unimelb.edu.au`)

◊ Head tutor: Matthew Farrugia-Roberts (`farrugiam1@unimelb.edu.au`)

◊ Tutors: Justin Tan (`justin.tan@unimelb.edu.au`)
Chenyuan Zhang (`chenyuanz1@unimelb.edu.au`)
Jason Liu (`jason.liu1@unimelb.edu.au`)
Josh Nguyen (`jtuandungn@unimelb.edu.au`)
Rohan Hitchcock (`hitchcockr@unimelb.edu.au`)
Scarlett Qin (`sijqin@unimelb.edu.au`)
Yifei Wang (`yifwang2@unimelb.edu.au`)

About Me

- Research interests:
  Machine Learning
  Large-Scale Data Mining
  Cyber Security
  Time Series Analysis
  Data Privacy

- Industry research partners:
  AARNet
  Schneider Electric
  Eye Nose and Ear Hospital
  AEMO

- Homepage: http://people.eng.unimelb.edu.au/smonazam/

General Information

- Text: *Artificial Intelligence: A Modern Approach*,
  Stuart Russell & Peter Norvig, 3rd Edition, Prentice Hall, 2014
- Lecture slides available on LMS, lectures recorded on Lecture Capture
- Subject LMS discussion board for student discussion
- Workshops run in a one hour and start second week of semester

Prerequisites

- ♢ Subjects:
  COMP20003 Algorithms and Data Structures or
  COMP20007 Design of Algorithms
- ♢ Skills:
  Data structures & algorithms coding in Python
  (This subject does not include programming language tuition)
  Familiarity with formal mathematical notation
  Basic understanding of differential calculus and probability theory helpful but not essential

Assessment

- ♢ Assessment: 70% exam, 30% project (programming project in Python)
- ♢ Requirements: 15/30 project hurdle, 35/70 exam hurdle, 50/100 overall
- ♢ Project: a single project in 2 parts
  Part A due 31st March. Part B due 12th May.
  (to be confirmed in project specification on subject LMS site)
- ♢ Project is to implement a game playing agent in Python
- ♢ You will work on the project in a team of two people
- ♢ We will discuss the project in more detail next lecture,
  and over the coming weeks

Who and Where

- Lectures:
  Tuesdays, 12–1 pm
  Thursdays, 12–1 pm

- Tutorials: (per your registration)

- Feedback:
  During/after lecture
  Tutorial
  Assignment feedback
  Discussion board
  Consultation sessions
  Matt, Thursdays 5–6 pm
  Sarah/Chris, Thursdays 1–1.30 pm
  General inquiries: comp30024-team@unimelb.edu.au

Syllabus

Topic	AIMA 2nd ed	AIMA 3rd ed
What is AI? (wk1)	Ch1	Ch1
Intelligent Agents (wk1)	Ch2	Ch2
Solving Problems by Searching (wk2)	Ch3	Ch3
Informed Search Methods (wk3)	Ch4	Ch3
Adversarial Search (wk4)	Ch6	Ch5
Learning in Games (wk5)	notes	notes
Feedback Quiz (wk6)	-	-
Vulnerabilities of AI (wk6)	-	-
Constraint Satisfaction Problems (wk7)	Ch5	Ch7
Uncertainty (wk8)	Ch13	Ch13
Probabilistic Reasoning (wk9)	Ch14	Ch14
Making Complex Decisions (wk10)	Ch17	Ch17
Robotics (wk11)	Ch25	Ch25
Revision, and Tournament (wk12)	-	-

Week 1: What is AI?

Chapter 1

Outline

◇ Defining AI

◇ Tests for intelligence

◇ State of the art

Types of Intelligence

◊ The big question: How does the mind arise from the brain?

◊ How many different types of "intelligent" behaviour can you think of?

Four approaches to defining AI

◊ Thinking like a human

◊ Thinking rationally

◊ Acting like a human

◊ Acting rationally

Thinking like a human

Cognitive modelling: figure out how we think by *introspection* or *experimentation*

Self-awareness is important: "I think therefore I am"

Humans feel emotions and apparently don't always think (or act) rationally

Thinking rationally

The laws of thought:
eg "Socrates is a man. All men are mortal.
Therefore Socrates is mortal"

Codifying rational thinking started with Aristotle (at least in the West)

The study of *logic* has greatly influenced AI

can't cope with new situation , complex scenario

Aspects of Intelligence

Abstract thinking and problem solving

Learning and memory

Language, communication

Intuition and creativity

Consciousness

Emotions

Surviving in a complex world

Adapting to new situations

Test for Intelligence: The Turing test

Proposed by Alan Turing in 1950

A human interrogates/converses with the computer via a teletype

The aim is for the computer to *imitate* a human well enough to fool people

Turing thought a computer would pass this test by the end of the century

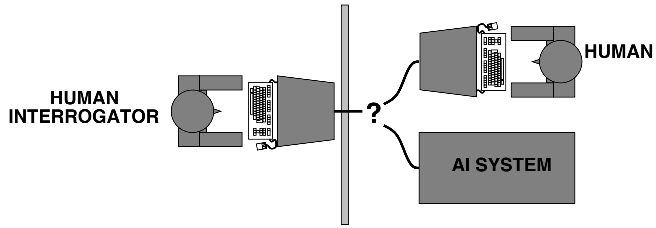
The Loebner Prize (Turing test) competition is held each year

The "total Turing test" allows physical objects to be passed to the machines as well as characters via the teletype

Acting humanly: The Turing test

Turing (1950) "Computing machinery and intelligence":

- ◇ "Can machines think?" ⟶ "Can machines behave intelligently?"
- ◇ Operational test for intelligent behavior: the Imitation Game



- ◇ Predicted that by 2000, a machine might have a 30% chance of fooling a lay person for 5 minutes
- ◇ Anticipated all major arguments against AI in following 50 years
- ◇ Suggested major components of AI: knowledge, reasoning, language understanding, learning

cannot repeat test

just tell pass or fail. doesn't tell improvement

Problem: Turing test is not *reproducible*, *constructive*, or amenable to *mathematical analysis*

very experimental

Eliza Session

The earliest attempt was a system called Eliza (1966).

Eliza: What brought you here to see me today?

Visitor: Oh, nothing much. I'm feeling a bit tired, that's all.

Eliza: Why do you think you're feeling a bit tired?

Visitor: Well, I've been traveling a lot, and away from home.

Eliza: Tell me more about your family

Eliza Session

Visitor: The trouble is, my mother's ill.

Eliza: How long has she been ill?

...

Visitor: The trouble is, my mother's Irish.

Eliza: How long has she been Irish?

...

Visitor: The trouble is, my mother's poodle.

Eliza: How long has she been poodle?

Acting rationally

The rational agent: perform actions which will (most likely) achieve one's *goals*

Knowledge may not be perfect — we need to go beyond strict rational thought in general

The rational agent view is the basis of "Artificial Intelligence: A Modern Approach"

State of the art

Which of the following can be done at present?

- ♢ Play a decent game of table tennis
- ♢ Drive along a curving mountain road
- ♢ Drive down Brunswick St on a Saturday night
- ♢ Play a decent game of bridge
- ♢ Discover and prove a new mathematical theorem
- ♢ Write an intentionally funny story
- ♢ Give competent legal advice in a specialized area of law
- ♢ Translate spoken English into spoken Japanese in real time

State of the Art

Machine translation: try Google Translator
(`https://translate.google.com`)

Conversational agents: Apple's Siri, IBM's Watson for question answering

Robotic vehicles: Google self-driving car autonomous vehicle that can drive safely though traffic
(`https://www.google.com/selfdrivingcar/`)

Versatile robots: 2015 DARPA Robotics Challenge - mobile robot that can walk over rubble and operate power tools

Human action recognition: Microsoft Kinect

Summary

◇ Defining AI
– Explain different approaches to defining AI

◇ Tests for intelligence
– Describe the operation of the Turing test

◇ State of the art
– Characterise the difficulty of different common tasks

What to do now:
– Find a project partner
– Brush up your Python
– Tutorials start in Week 2

Week 1: Intelligent Agents

Chapter 2

Outline

◊ Agent model

◊ Agent types

◊ Environment types

◊ Summary

Intelligent agents

◊ chess/backgammon

◊ refinery controller

◊ medical diagnosis

◊ flight reservations

◊ walking on two legs

◊ taxi driver

◊ vacuum cleaning

◊ robocup soccer

The Agent Model

- ♢ **Percepts**/observations of the environment, made by *sensors*
- ♢ **Actions** which may affect the environment, made by *actuators*
- ♢ **Environment** in which the agent exists
- ♢ **Performance measure** of the desirability of environment states

Example: automated taxi

Percepts?? video, accelerometers, gauges, engine sensors, keyboard, GPS, ...

Actions?? steer, accelerate, brake, horn, speak/display, ...

Environment?? city streets, freeways, traffic, pedestrians, weather, customers, ...

Performance measure?? safety, reach destination, maximize profits, obey laws, passenger comfort, ...

Agents as functions

Agents can be evaluated empirically, sometimes analysed mathematically

Agent is a function from *percept sequences* to actions

Ideal rational agent would pick actions which are expected to maximise its *performance measure* (based on the percept sequence and its built-in knowledge)

Rational $\neq$ omniscient   no guarantee for the outcome
Rational $\neq$ clairvoyant
Rational $\neq$ successful

Agent types

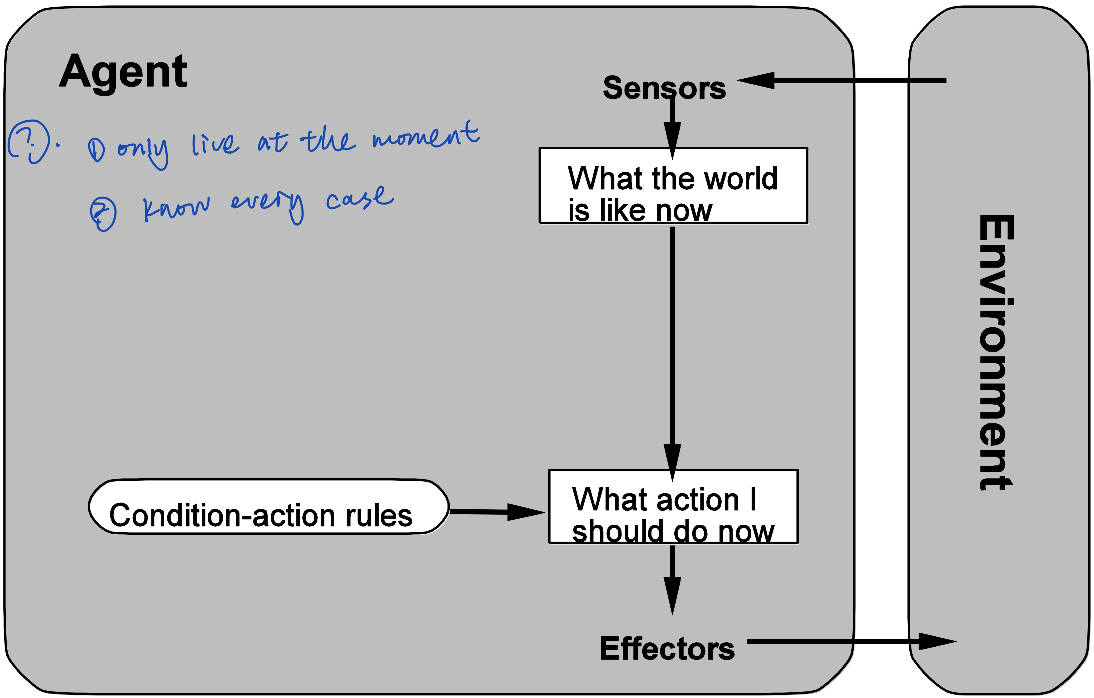
◊ simple reflex agents

◊ model-based reflex agents

◊ goal-based agents

◊ utility-based agents

Simple reflex agents

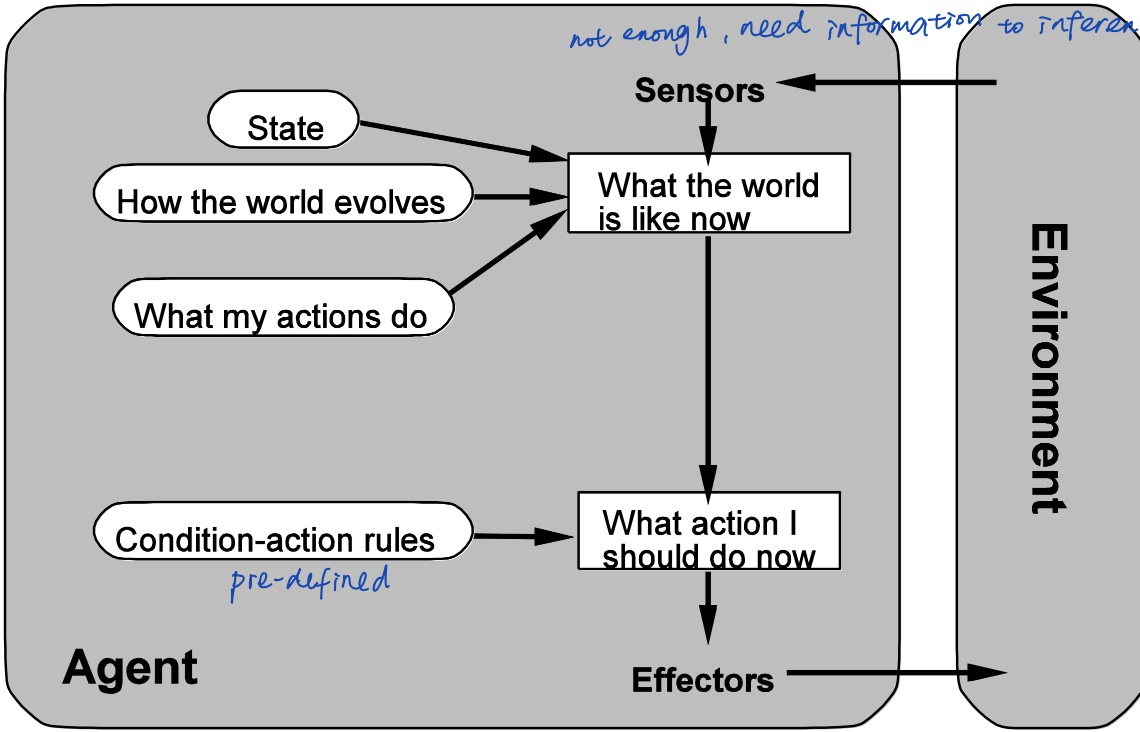


Model-based reflex agents

cannot plan

assume rule not fully observable, remember previous information

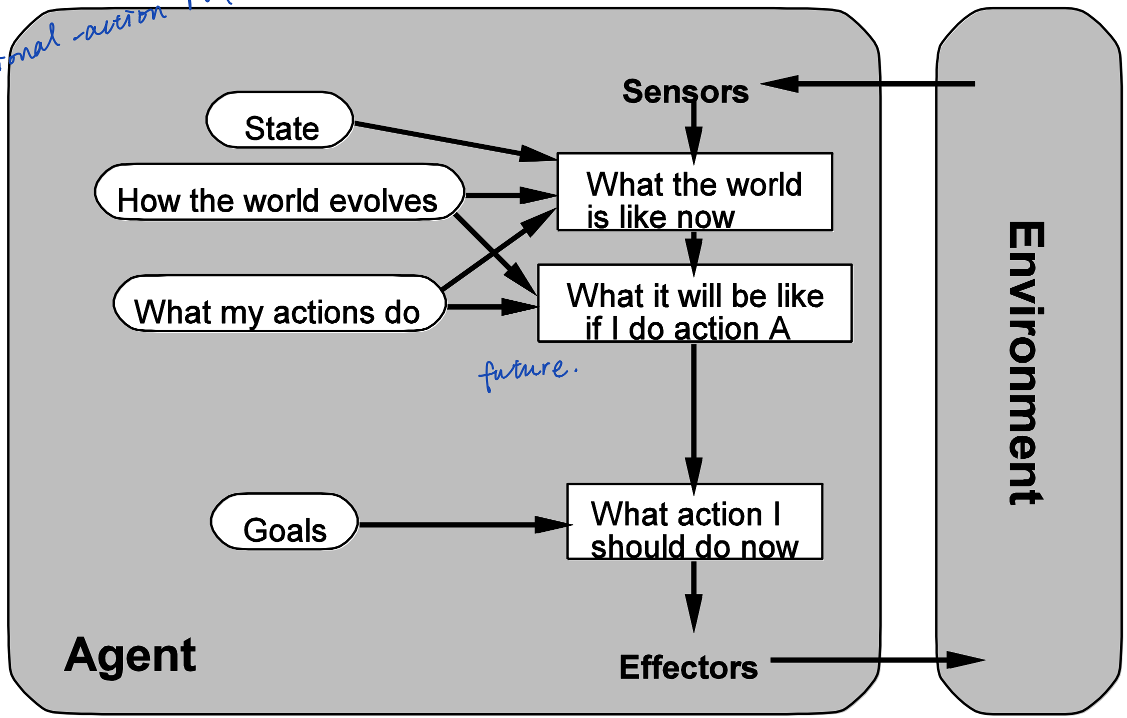
not enough, need information to inference.



pre-defined

Goal-based agents

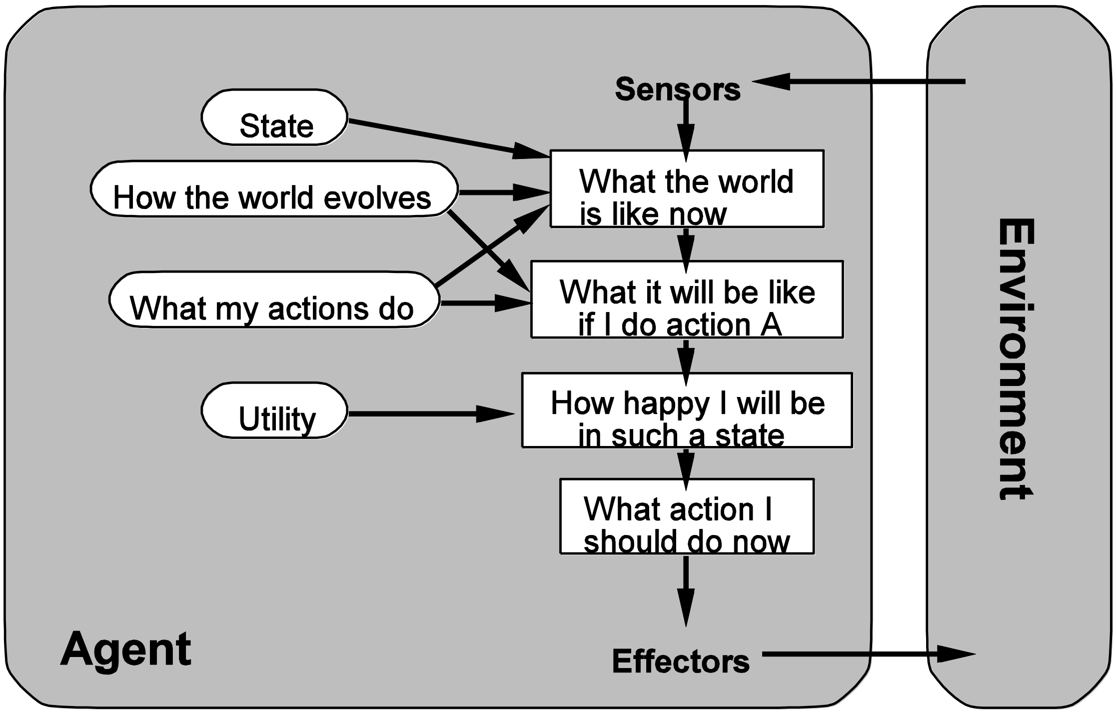
not use conditional-action rule



future.

Utility-based agents

prioritize goals. satisfy as many goals as possible



Environment types

Environments may or may not be

{ fully observable
  partial observable

◇ **Observable**: percept contains all relevant information about the world

{ deterministic
  stochastic

◇ **Deterministic**: current state of the world uniquely determines the next

{ episodic
  sequential

◇ **Episodic**: only the current (or recent) percept is relevant ⇒ simple reflex agent

{ static
  dynamic

◇ **Static**: environment doesn't change while the agent is deliberating

◇ **Discrete**: finite number of possible percepts/actions

Environment types

	Solitaire	Backgammon	Internet shopping	Taxi
Observable??				
Deterministic??				
Episodic??				
Static??				
Discrete??				

The environment type largely determines the agent design

The real world is (of course) partially-observable, stochastic, sequential, dynamic, continuous

Environment types

	Solitaire	Backgammon	Internet shopping	Taxi
Observable??	Yes	Yes	No	No
Deterministic??	Yes	No	Partly	No
Episodic??	No	No	No	No
Static??	Yes	Yes	Semi	No
Discrete??	Yes	Yes	Yes	No

The environment type largely determines the agent design

The real world is (of course) partially-observable, stochastic, sequential, dynamic, continuous

Summary

◇ Agent model
– characterise requirements for an agent in terms of its percepts, actions, environment and performance measure

◇ Agent types
– choose and justify choice of agent type for a given problem

◇ Environment types
– characterise the environment for a given problem