

COMP250: Artificial Intelligence

2: Designing AI behaviours

Reminder

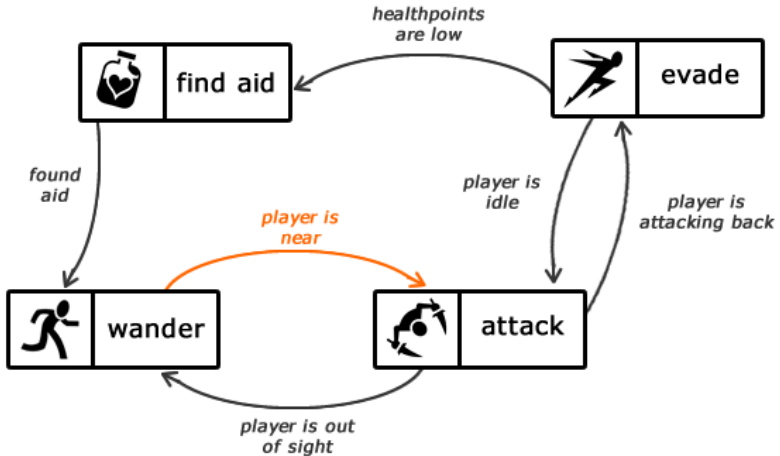
AI component proposal due **next week!**

AI architectures

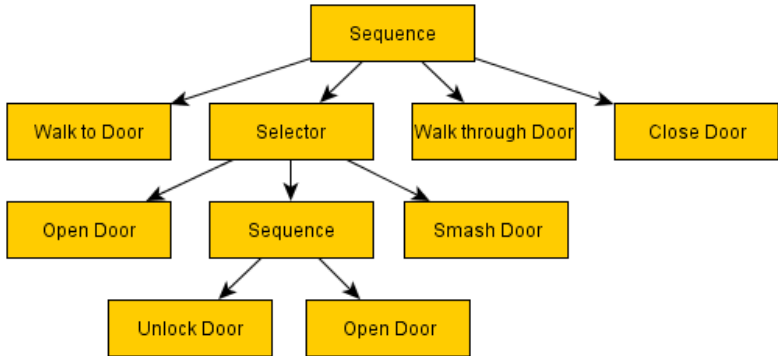
Rule-based AI

Generally implemented as `if` statements or event-based triggers

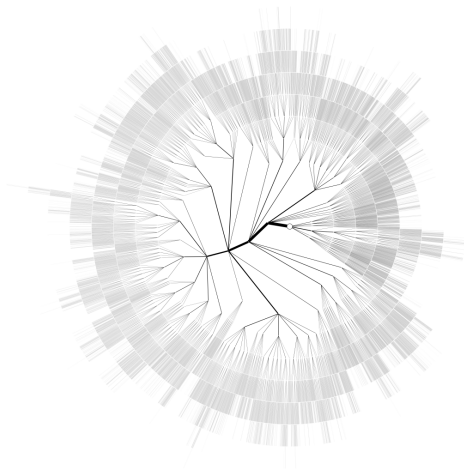
Finite state machines



Behaviour trees



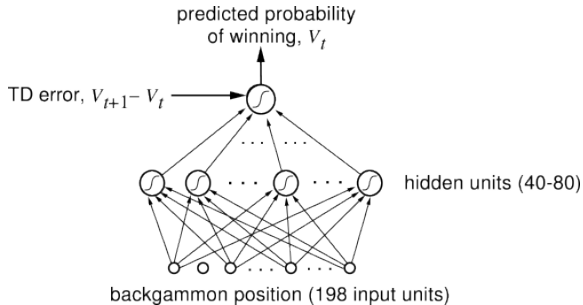
Game tree search



Multi-agent approaches (e.g. flocking)



Machine learning



AI architectures

- ▶ Can roughly be divided into **hand-authored**...
 - ▶ Rule-based, FSM, behaviour trees
- ▶ ... and **computational intelligence**
 - ▶ Search, multi-agent, machine learning
- ▶ Do you want to **design** the AI behaviours yourself, or do you want them to **emerge** from the system?
- ▶ Predictability and authorial control versus unpredictability and novelty

Logic

Logical operations

Python	C family	Mathematics	Behaviour tree
not a a and b a or b	!a a && b a b	$\neg A$ or \bar{A} $A \wedge B$ $A \vee B$	Inverter Sequence Selector

The laws of thought

- ▶ Let A be a **proposition** (a statement about the world)
- ▶ A is a **boolean** value, either **true** or **false**
- ▶ The law of **identity**: $A == A$ is always true
- ▶ The law of **non-contradiction**: $A \ \&\& \ !A$ is always false
 - ▶ I.e. A cannot be both true and false
- ▶ The law of the **excluded middle**: $A \ || \ !A$ is always true;
 - ▶ I.e. A must be either true or false

Predicates

- ▶ **Predicates** are propositions with **parameters**
- ▶ In programming terms, a predicate is a function that returns a boolean
- ▶ E.g. `LivesIn(Bob, Falmouth)` could be a predicate for "Bob lives in Falmouth"

Quantifiers

- ▶ $P(x)$ is a predicate
- ▶ $\forall x : P(x)$ means that $P(x)$ is true **for all** values of x
- ▶ $\exists x : P(x)$ means that **there exists** at least one value of x such that $P(x)$ is true

Implication

- ▶ “ A implies B ” means “if A is true then B is true”
- ▶ Written as $A \implies B$
- ▶ E.g. if someone lives in Falmouth then they live in Cornwall
- ▶ $\forall x : \text{LivesIn}(x, \text{Falmouth}) \implies \text{LivesIn}(x, \text{Cornwall})$

Contrapositive

- ▶ $A \implies B$ is equivalent to $\neg B \implies \neg A$
- ▶ E.g. if someone does not live in Cornwall then we know they don't live in Falmouth
- ▶ $\forall x : \neg \text{LivesIn}(x, \text{Cornwall}) \implies \neg \text{LivesIn}(x, \text{Falmouth})$

Equivalence

- ▶ If $A \implies B$ and $B \implies A$ then A and B are **logically equivalent**
- ▶ A is true **if and only if** B is true
- ▶ Written as $A \iff B$
- ▶ E.g. “Alice lives in a city in Cornwall” if and only if “Alice lives in Truro”
- ▶ This relies on an extra piece of domain knowledge: Truro is the only city in Cornwall
 - ▶ $\forall x : \text{InCornwall}(x) \wedge \text{IsCity}(x) \implies x = \text{Truro}$

Implication is transitive

- ▶ If $A \implies B$ and $B \implies C$ then $A \implies C$
- ▶ E.g. if someone lives in Falmouth then they live in Cornwall
- ▶ And if someone lives in Cornwall then they live in England
- ▶ Therefore if someone lives in Falmouth then they live in England

Inverting quantifiers

- ▶ “Everyone who lives in Cornwall likes cider”
- ▶ $\forall x : \text{LivesIn}(x, \text{Cornwall}) \implies \text{Likes}(x, \text{Cider})$
- ▶ What is the **opposite** of this statement?
- ▶ $\neg(\forall x : \text{LivesIn}(x, \text{Cornwall}) \implies \text{Likes}(x, \text{Cider}))$
- ▶ In logical terms, the opposite is **not** “nobody who lives in Cornwall likes cider”
- ▶ It’s “Not everyone who lives in Cornwall likes cider”
- ▶ I.e. “There is at least one person living in Cornwall who does not like cider”
- ▶ $\exists x : \text{LivesIn}(x, \text{Cornwall}) \wedge \neg \text{Likes}(x, \text{Cider})$

Game theory

Game theory

- ▶ A branch of mathematics studying **decision making**
- ▶ A **game** is a system where one or more **players** choose **actions**; the combination of these choices lead to each agent receiving a **payoff**
- ▶ Important applications in economics, ecology and social sciences as well as AI

The Prisoner's Student's Dilemma

- ▶ Two students, **Alice** and **Bob**, are suspected of copying from each other
- ▶ Each is offered a deal in exchange for information
- ▶ Each can choose to **betray** the other or stay **silent** — but they **cannot communicate** before deciding what to do
- ▶ If **both stay silent**, both receive a C grade
- ▶ If **Alice betrays Bob**, she receives an A whilst he gets expelled
- ▶ If **Bob betrays Alice**, he receives an A whilst she gets expelled
- ▶ If **both betray each other**, both get an F

Payoff matrix

	A silent	A betray
B silent	A: 50 B: 50	A: 70 B: -100
B betray	A: -100 B: 70	A: 0 B: 0

Socratic FALCOMPED: what would you do?

Nash equilibrium

- ▶ Consider the situation where both have chosen to betray
- ▶ Neither person has anything to gain by switching to silence, assuming the other person doesn't also switch
- ▶ Such a situation is called a **Nash equilibrium**
- ▶ If all players are **rational** (in the sense of wanting to maximising payoff), they should converge upon a Nash equilibrium

Does every game have a Nash equilibrium?

	A rock	A paper	A scissors
B rock	A: 0 B: 0	A: +1 B: -1	A: -1 B: +1
B paper	A: -1 B: +1	A: 0 B: 0	A: +1 B: -1
B scissors	A: +1 B: -1	A: -1 B: +1	A: 0 B: 0

Socratic FALCOMPED: what would you do?

Nash equilibrium for Rock-Paper-Scissors

- ▶ Committing to any choice of action can be **exploited**
- ▶ E.g. if you always choose paper, I choose scissors
- ▶ If we try to reason naïvely, we get stuck in a loop
 - ▶ If I choose paper, you'll choose scissors, so I should choose rock, but then you'll choose paper, so I'll choose scissors, so you'll choose rock, so I choose paper...
- ▶ The optimum strategy is to be **unpredictable**
- ▶ Choose rock with probability $\frac{1}{3}$, paper with probability $\frac{1}{3}$, scissors with probability $\frac{1}{3}$

Mixed strategies

- ▶ A **mixed strategy** assigns probabilities to actions and chooses one at random
- ▶ In contrast to a **pure** or **deterministic strategy**, which always chooses the same action
- ▶ If we allow mixed strategies, **every game has at least one Nash equilibrium**

Guess $\frac{2}{3}$ of the average

- ▶ Everyone guesses a real number (decimals are allowed) between 0 and 100 inclusive
- ▶ The winner is the person who guesses closest to $\frac{2}{3}$ of the mean of all guesses
- ▶ Example:
 - ▶ If the guesses are 30, 40 and 80...
 - ▶ ... then the mean is $\frac{30+40+80}{3} = 50...$
 - ▶ ... so the winning guess is 30, as this is closest to $\frac{2}{3} \times 50 = 3.333$
- ▶ Socrative FALCOMPED: make your guesses!

Rationality

- ▶ Rationality is a useful assumption for mathematics and AI programmers
- ▶ However it's important to remember that **humans aren't always rational**