

# CSU33061 Artificial Intelligence I

## Introduction

`www.scss.tcd.ie/Tim.Fernando/AI`

### *Key Phrases:*

*Can machines think?*

- Turing test & ELIZA effect
- AI-complete (contra low hanging fruit)

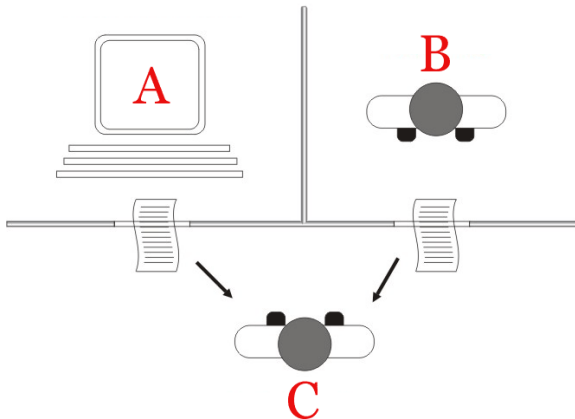
Agent & environment

- Cognitive Revolution & Big Data

Levels of intelligence

# Can machines think? (Turing 1950)

**Turing test:** can C tell A from B?

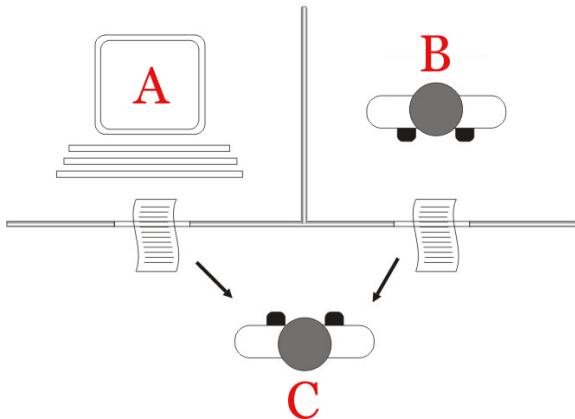


From Wikipedia, (Juan Alberto Sánchez Margallo)

Intelligence operationalized: subject to testing

# Can machines think? (Turing 1950)

**Turing test:** can C tell A from B?



From Wikipedia, (Juan Alberto Sánchez Margallo)

Intelligence operationalized: subject to testing  
... cheating?

# ELIZA (Weizenbaum, 1964-66) & artful deception

- use pattern matching and substitution to fake understanding

## ELIZA (Weizenbaum, 1964-66) & artful deception

- use pattern matching and substitution to fake understanding

**ELIZA effect:** humans are inclined to see computers as humans

e.g. when ATM says “thank you”

## ELIZA (Weizenbaum, 1964-66) & artful deception

- use pattern matching and substitution to fake understanding

**ELIZA effect:** humans are inclined to see computers as humans

e.g. when ATM says “thank you”

An AI problem is **AI-complete** if any AI problem is mechanically reducible to it (i.e., it is at least as hard as any other).

E.g. Natural Language Understanding

*The town councilors refused to give the demonstrators a permit because **they** feared violence.*

*Who feared violence?*

T. Winograd

## ELIZA (Weizenbaum, 1964-66) & artful deception

- use pattern matching and substitution to fake understanding

**ELIZA effect:** humans are inclined to see computers as humans

e.g. when ATM says “thank you”

An AI problem is **AI-complete** if any AI problem is mechanically reducible to it (i.e., it is at least as hard as any other).

E.g. Natural Language Understanding

*The town councilors refused to give the demonstrators a permit because **they** advocated violence.*

*Who advocated violence?*

T. Winograd

## ELIZA (Weizenbaum, 1964-66) & artful deception

- use pattern matching and substitution to fake understanding

**ELIZA effect:** humans are inclined to see computers as humans

e.g. when ATM says “thank you”

An AI problem is **AI-complete** if any AI problem is mechanically reducible to it (i.e., it is at least as hard as any other).

E.g. Natural Language Understanding

*The town councilors refused to give the demonstrators a permit because **they** advocated violence.*

*Who advocated violence?*

T. Winograd

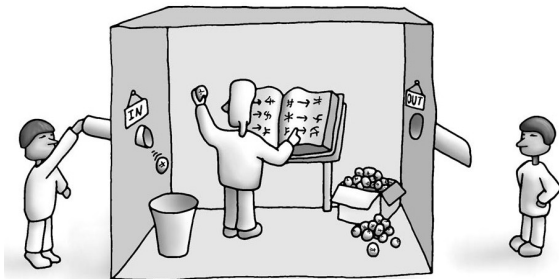
CAUTION: Programs may appear to work better than they do

**Siri rage** (Urban dictionary):

*When you get enraged because Siri just doesn't get it.*



## Chinese room argument (Searle's thought experiment)

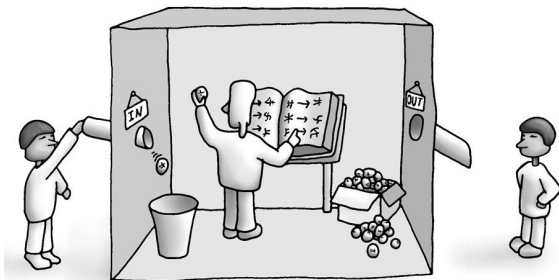


From <http://america.pink/images/9/6/3/2/5/4/en/2-chinese-room.jpg>

- a clerk can follow instructions for communicating in Chinese without understanding Chinese

# Strong vs Weak AI

## Chinese room argument (Searle's thought experiment)



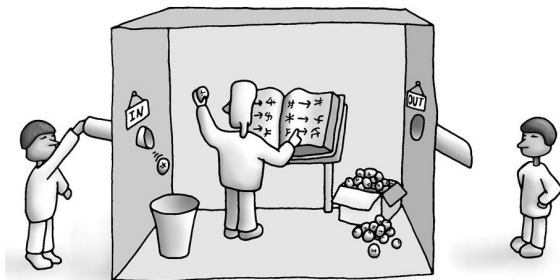
From <http://america.pink/images/9/6/3/2/5/4/en/2-chinese-room.jpg>

- a clerk can follow instructions for communicating in Chinese without understanding Chinese

mindless obedience  $\neq$  mind from biological processes

# Strong vs Weak AI

## Chinese room argument (Searle's thought experiment)



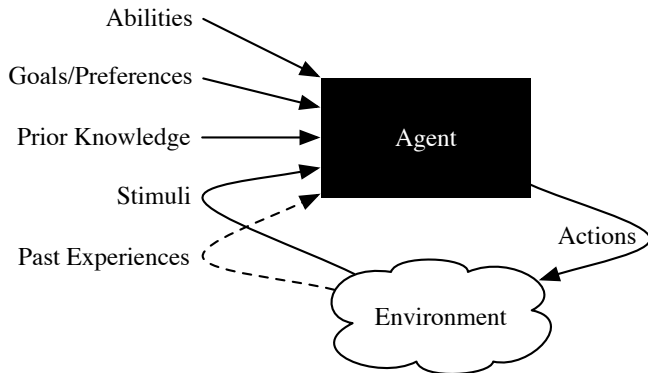
From <http://america.pink/images/9/6/3/2/5/4/en/2-chinese-room.jpg>

- a clerk can follow instructions for communicating in Chinese without understanding Chinese

mindless obedience  $\neq$  mind from biological processes

**Just do it**  $\rightsquigarrow$  Black Box, judged by its actions

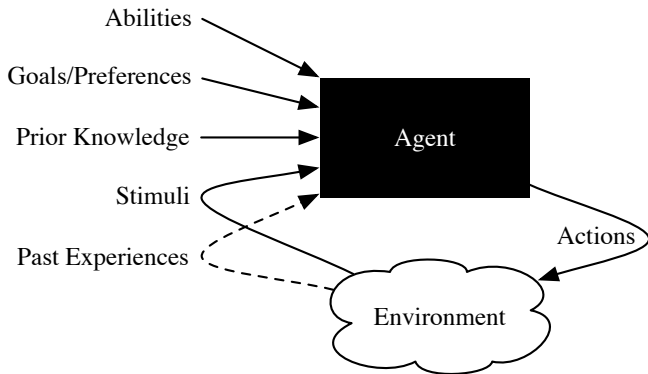
# Locating intelligence (black box)



Poole & Mackworth

Intelligence: (abilities, goals, ..., experience)  $\mapsto$  action

## Locating intelligence (black box)



Poole & Mackworth

Intelligence: (abilities, goals, ..., experience)  $\mapsto$  action

Turing test: what to say  $\rightsquigarrow$  what to do

## Between agent and environment

agent	environment
program	data
Cognitive Revolution	BIG DATA
hard-wired	experienced
rationalist	empiricist
nativist	behaviorist
innate	tabula rasa
nature	nurture

Turing machine &  
specialized automaton

## Between agent and environment

agent	environment
program	data
Cognitive Revolution	BIG DATA
hard-wired	experienced
rationalist	empiricist
nativist	behaviorist
innate	tabula rasa
nature	nurture

Turing machine &  
specialized automaton

Learning (from environment)  
trial & error: “data as oil”

## Between agent and environment

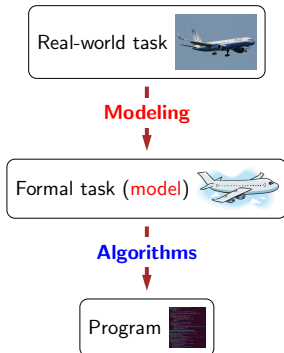
agent	environment
program	data
Cognitive Revolution	BIG DATA
hard-wired	experienced
rationalist	empiricist
nativist	behaviorist
innate	tabula rasa
nature	nurture

Turing machine &  
specialized automaton

Learning (from environment)  
trial & error: “data as oil”

Moving target: changing agent & environment  
e.g. change in state



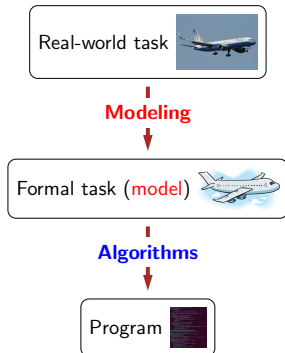


unstructured information  $\rightsquigarrow$  actionable knowledge

Demis Hassabis

[www.theguardian.com/technology/2016/feb/16/demis-hassabis-artificial-intelligence-deepmind-alphago](http://www.theguardian.com/technology/2016/feb/16/demis-hassabis-artificial-intelligence-deepmind-alphago)

# From Narrow to General AI



unstructured information  $\rightsquigarrow$  actionable knowledge

Demis Hassabis

[www.theguardian.com/technology/2016/feb/16/demis-hassabis-artificial-intelligence-deepmind-alphago](http://www.theguardian.com/technology/2016/feb/16/demis-hassabis-artificial-intelligence-deepmind-alphago)

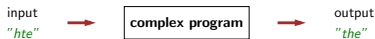
Autonomous = perform tasks in complex environments without  
constant user guidance

Adaptive = improve performance by learning from experience



## Traditional approach

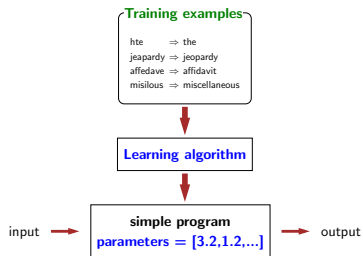
A spell checker:



Problem: complexity becomes unwieldy



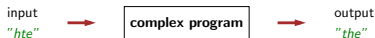
## Machine learning approach





## Traditional approach

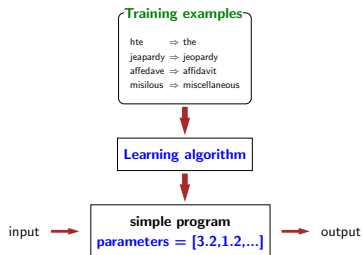
A spell checker:



Problem: complexity becomes unwieldy



## Machine learning approach



Search problems

Markov decision processes

Adversarial games

Constraint satisfaction problems

Bayesian networks

**Reflex**

**States**

**Variables**

**Logic**

"Low-level intelligence"

"High-level intelligence"

**Machine learning**

# Back in Trinity

## Undergraduate ML modules

- ▶ CSU44061 Machine Learning  
Semester 1 (5 ECTS)
- ▶ CSU44062 Advanced Computational Linguistics  
Semester 1 (5 ECTS)  
unsupervised ML for natural language processing

# Back in Trinity

## Undergraduate ML modules

- ▶ CSU44061 Machine Learning  
Semester 1 (5 ECTS)
- ▶ CSU44062 Advanced Computational Linguistics  
Semester 1 (5 ECTS)  
unsupervised ML for natural language processing

## CSU33061: a taste building on CSU34011 (Prolog)

- ▶ logic & agents as Turing machines
- ▶ search
- ▶ Q-learning & Markov decision processes
- ▶ Constraint satisfaction
- ▶ Bayesian & Markov networks