CS4060 Knowledge Representation & Automata

Introduction

www.scss.tcd.ie/Tim.Fernando/KRA

2017-18

CS4404 Machine Learning
CS4LL5 Unsupervised ML for NLP (since 2014)
CS4060 Knowledge Representation & Automata

Introduction

www.scss.tcd.ie/Tim.Fernando/KRA

2017-18

CS4404 Machine Learning
CS4LL5 Unsupervised ML for NLP (since 2014)

2014-2017

CS4061 Artificial Intelligence (Michaelmas)
CS4062 Machine Learning for NLP (Hilary)
CS4060 Knowledge Representation & Automata

Introduction

www.scss.tcd.ie/Tim.Fernando/KRA

2017-18

CS4404 Machine Learning
CS4LL5 Unsupervised ML for NLP (since 2014)

2014-2017

CS4061 Artificial Intelligence (Michaelmas)
CS4062 Machine Learning for NLP (Hilary)

1999-2015

CS4LL4ab Artificial Intelligence (year-long)

- logical orientation (Prolog, over LISP)
  Logicomix, by Doxiadis & Papadimitriou
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 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 (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 (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 (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.
Locating intelligence (black box)

Intelligence: \((\text{abilities, goals, \ldots, experience}) \mapsto \text{action}\)

Poole & Mackworth
Locating intelligence (black box)

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

Turing test: what to say \( \rightsquigarrow \) what to do
## Between agent and environment

<table>
<thead>
<tr>
<th>agent</th>
<th>environment</th>
</tr>
</thead>
<tbody>
<tr>
<td>program</td>
<td>data</td>
</tr>
<tr>
<td>Cognitive Revolution</td>
<td>Big Data</td>
</tr>
<tr>
<td>hard-wired</td>
<td>experienced</td>
</tr>
<tr>
<td>rationalist</td>
<td>empiricist</td>
</tr>
<tr>
<td>nativist</td>
<td>behaviorist</td>
</tr>
<tr>
<td>innate</td>
<td>tabula rasa</td>
</tr>
<tr>
<td>nature</td>
<td>nurture</td>
</tr>
</tbody>
</table>

Turing machine &
specialized automaton

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

<table>
<thead>
<tr>
<th>agent</th>
<th>environment</th>
</tr>
</thead>
<tbody>
<tr>
<td>program</td>
<td>data</td>
</tr>
<tr>
<td>Cognitive Revolution</td>
<td>Big Data</td>
</tr>
<tr>
<td>hard-wired</td>
<td>experienced</td>
</tr>
<tr>
<td>rationalist</td>
<td>empiricist</td>
</tr>
<tr>
<td>nativist</td>
<td>behaviorist</td>
</tr>
<tr>
<td>innate</td>
<td>tabula rasa</td>
</tr>
<tr>
<td>nature</td>
<td>nurture</td>
</tr>
</tbody>
</table>

Turing machine &
specialized automaton

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

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

Real-world task

Modeling

Formal task (model)

Algorithms

Program

unstructured information $\rightsquigarrow$ actionable knowledge

Demis Hassabis

www.theguardian.com/technology/2016/feb/16/demis-hassabis-artificial-intelligence-deepmind-alphago
Levels of intelligence

Search problems
Markov decision processes
Adversarial games
Constraint satisfaction problems
Bayesian networks

Reflex States Variables Logic

"Low-level intelligence" "High-level intelligence"

Machine learning

https://web.stanford.edu/class/cs221/ (Autumn 2017)
Levels of intelligence

Search problems
Markov decision processes
Adversarial games
Constraint satisfaction problems
Bayesian networks

"Low-level intelligence" "High-level intelligence"

Machine learning

https://web.stanford.edu/class/cs221/ (Autumn 2017)

A Pendulum Swung Too Far, Kenneth Church,
Linguistic Issues in Language Technology 2 (4), 2011

CS4060 focuses on the right side, logic

▶ what we can say (& can’t): knowledge representation
▶ what we can do (& can’t): automata
Resources
(links in www.scss.tcd.ie/Tim.Fernando/KRA)

- *AI: Foundations of Computational Agents*
  Poole & Mackworth, 2nd edition (2017)

- *Knowledge Representation & Reasoning*
  Brachman & Levesque

- *Elements of Finite Model Theory*, L. Libkin

- *Algorithmic Model Theory*, E. Grädel

- Predication via Finite-State Methods, ESSLLI 2017 course