Utrecht University

5-9-2019

Methods in AI Research

Introduction

Floris Bex
MAIR course coordinator

Floris Bex

- Professor of Data Science and the Judiciary
  - Law and Technology (Tilburg)
- Scientific Director Police Lab AI
  - Innovation Center for AI (ICAI – A'dam/Utrecht)
- Researcher & Lecturer Intelligent Systems
  - Information and Computing Sciences (Utrecht)

Floris Bex - research

- Computational argumentation
  - Formal (logical, probabilistic) models of argument and debate
    - Using argumentation to explain machine learning
    - Using argumentation to drive police intake dialogues
    - Using argumentation to reason about evidence in legal cases
- AI & Law, policing
  - AI for the legal field and the police
    - Making concept verdicts for judges, summarizing legal cases
    - Evidence analysis using Bayesian Networks, argument-based intake
  - Relation between AI, the law and policing
    - Explaining AI decision to judges, predictive policing, robojudges

Floris Bex - research

- Computer science perspective…
  - Engineering smart systems
- …with a little law and philosophy thrown in.
  - What are the legal and ethical implications of AI? Which type of AI fits the law best?

Methods in AI research

- AI has roots in various fields
  - Logic, mathematics, psychology, neurology, computer science, linguistics
- These fields have different languages and (research) methods
  - Statistics, logic, algorithms, natural language
  - Engineering, empirical evaluation, mathematical proofs

Methods in AI research – an intro to Utrecht AI…

Know and understand the basics of:

- Agents
  - Dialogue modelling, autonomous chatbot/dialogue system
- Reasoning
  - Knowledge-based systems, description & argumentation logics
- Cognitive processing
  - Cognitive modelling, experimentation, interaction
- Natural language
  - NLP, interaction
- and of course machine learning!

Methods in AI research – …and its methods

Be able to use different research methods in AI, such as:

- Implementing a working AI system
- Testing and evaluating an AI system (technical capabilities, performance, usability)
- Writing a technical report and a research paper on an AI system, its evaluation and its place in the broader context of AI

Methods in AI research – …and its methods

- Building, evaluating and writing about an autonomous system that understands natural language, interacts with the user in reasoned dialogue
  - Understanding and modeling a specific knowledge domain
  - Implementing and empirically evaluating a machine learning-based NLP algorithm
  - Implement a working AI system using Python
  - Designing, conducting an experiment with human participants to empirically evaluate your system
  - Writing technical/scientific reports about your system, its relation to other systems in the literature, and the empirical evaluation

What is AI?

- What is AI research? → What is AI?

What is AI?

- Accurately being able to distinguish cats from dogs?
- Winning at chess?
- Determining the amount of tax to be paid?
- Winning at Go?
- Driving a car?
- Predicting where the next burglary will likely be?
- Answering questions asked in natural language?
- Passing the Turing test?

One (wrong) definition of AI

"The AI in question, machine learning, is a technique for recognising patterns in relevant and preferably as complete as possible data files with the aim of discovering patterns in reality."

Minister of Justice to Parliament of the Netherlands

One (right) definition of AI

- Making computers perform tasks for which intelligence is required if performed by humans
  - Sense the environment;
  - Reason with knowledge;
  - Plan actions;
  - Learn new knowledge;
  - Communicate and interact with humans and other AI;
  - and integrate all these skills to achieve goals

Another (right) definition of AI

"Systems that exhibit intelligent behaviour by analysing their environment and - with a certain degree of autonomy - taking action to achieve specific objectives."

European Commission
Coordinated strategy on AI

Strong AI vs Weak AI

- An artificial intelligence system can think and have a mind.
  - But how do we know the system is thinking?
  - An artificial intelligence system can act like it thinks and has a mind or simulate thinking
- The Chinese Room (Searle)

Strong AI vs. Weak AI

- Strong AI: machines that can think, experience emotions, consciousness etc.
- Weak AI: machines that can perform tasks for which some intelligence is needed

- As soon as a part of AI is successful, it becomes a research area
  - advanced search, language understanding, machine learning etc. etc.

"As soon as it works, no one calls it AI any more."

John McCarthy

Strong AI vs Weak AI – the computer science perspective

"Most AI researchers take the weak AI hypothesis for granted, and don't care about the strong AI hypothesis—as long as their program works, they don't care whether you call it a simulation of intelligence or real intelligence." *Russel & Norvig, AI – a modern approach, 2009*

"The question of whether Machines Can Think... ...is about as relevant as the question whether Submarines Can Swim" *Edsger Dijkstra, The threats to computing science, 1984*

Views on AI

Thinking humanly **The cognitive modeling approach**	Thinking rationally **The "laws of thought" approach**
Acting humanly **The Turing Test approach**	Acting rationally **The rational agent approach**

Acting humanly – the Turing Test approach

- Turing Test – does a human know they're talking to a computer?
  - Coffee Test, Robot College Student Test, Employment Test, Flat Pack Furniture Test
  - Conversation with Eliza, Google Duplex
- Suggested all major components of AI: knowledge, reasoning, language understanding, learning, (vision, robotics)
- Affective computing, care robots

Thinking humanly – the Cognitive Modelling Approach

- How do humans think?
  - Predicting and testing behaviour of human subjects (top-down, cognitive science)
  - Direct identification from neurological data (bottom-up, cognitive neuroscience)
- How can we simulate this in computers?
  - Cognitive architectures, neural networks
- Why simulate?
  - To know more about human intelligence
  - To have more natural AI, explainable AI

Thinking rationally – the "laws of thought" approach

- Logic: notation and rules for correct reasoning or arguments
  - Classical logic, deontic logic, temporal logic, fuzzy logic, epistemic logic, probabilistic logic, …
- Direct line through mathematics and philosophy to modern (symbolic) AI
- Very useful for more formal reasoning
  - E.g. normative or legal reasoning
- Problems of complexity, dealing with the real world
  - Combination with sub-symbolic (machine learning)

Acting rationally – the Rational Agent approach

- Rational behaviour: doing the thing, which is expected to maximize goal achievement, given the available information

- More general than “laws of thought”
  - Not just correct inference
- Easier to measure what “the right thing” is than for the cognitive modelling approach
  - Rational standards
- Negotiation, game theory

Views on AI in Utrecht

- Acting humanly – natural language, cognitive processing
- Thinking humanly – cognitive processing, natural language
- Thinking rationally – reasoning, agents
- Acting rationally – agents, reasoning

All of them are useful and learn from each other in modern AI

AI – science or engineering?

- Engineering ("building systems") always a big part of AI
  - current AI hype is because of working systems
- Many AI researchers care less about engineering real systems, more about theory
  - plus, resources needed to build systems are large
- Theories are needed to approach engineering problems
  - Engineering is needed to answer scientific questions
- In Utrecht: most engineering in Agents, but also in Cognitive Processing and Natural Language

What kind of AI are these well-known projects?

- Deep Blue Chess
- Watson Jeopardy
- Self driving vehicles
- Google AlphaGo
- Google AlphaZero
- Conversation with Eliza
- Google Duplex

MAIR topics & lectures

- Dialogue systems
- Machine Learning
- Natural Language Processing
- Linguistic & social interactions
- Cognitive Modelling
- Experimentation
- Scientific writing
- Designing responsible AI
- Knowledge-based systems
- Logic & argumentation

Floris Bex

Dong Nguyen

Chris Janssen

Rosalie Iemhoff

Lecturers

- Different lecturers, different subfields of AI, different styles
- Lecturers will indicate clearly during the lectures what you should know and which skills you are expected to acquire
- This lecture: background (not exam material)

Floris Bex

Dong Nguyen

Chris Janssen

Rosalie Iemhoff

MAIR Project – restaurant recommendation dialogue system

- Building, evaluating and writing about an autonomous system that understands natural language, interacts with the user in reasoned dialogue

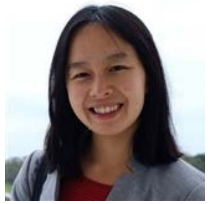
S: Hello, how may I help you?
U: I want an expensive place that serves English food.
S: Cote is a nice expensive restaurant with English food.
U: What is the phone number?
S: Its number is 01223 311053.
U: Thanks for the help, goodbye.
S: Thank you, goodbye!

S: System
U: User

MAIR Project – restaurant recommendation dialogue system

- Part 1 - designing and implementing dialogue model, implementing and evaluating machine learning classifier for natural language, implementing different interaction features.

- Part 2 - designing, carrying out and reporting on user experiments with your system, as well as thinking & writing about your system in the wider context of AI.



Dong Nguyen

Marijn Schraagen

Chris Janssen

Rosalie Iemhoff

MAIR project & lectures

- Dialogue systems
- Machine Learning
- Natural Language Processing
- Linguistic & social interactions
- Cognitive Modelling
- Experimentation
- Scientific writing
- Designing responsible AI
- Knowledge-based systems
- Logic & argumentation

MAIR labs

- Mixed teams of 4, made by us
- Register and tell us your skills via Google Form
  - **Deadline: Friday 6 September**
- Team and lab room assignment: Tuesday morning the 10th


Marijn Schraagen


Onuralp Ulusoy


Davide Dell'Anna


Marc Schwartz

MAIR communication

- Blackboard is the most important resource
  - Information on Blackboard or sent via Blackboard is “official”

- Questions about the assignment?
  - TA (use the lab sessions)
- General questions about the course?
  - Floris Bex
- Specific questions about a lecture?
  - Lecturer (use the lectures)

MAIR grading

- 30% project part 1
- 30% project part 2
- 40% individual final exam with theory-questions

- All three grades need to be at least 5.0 unrounded. Furthermore, the weighted final grade needs to be at least 5.5 unrounded.
- If you fail to pass one of the project parts then you can submit an improved version of that part of the project for a maximum grade of 5.5 for the resubmitted part. The deadline and requirements for improvements are set on an individual basis.
- If you fail to pass the first exam you can attend the retake exam in January.

MAIR and the rest of your studies

- We will take time (19 Sept) to explain the curriculum, areas and possible study paths
- Area representatives: Bex (agents), Nouwen (language), Janssen (cognitive processing), Iemhoff (reasoning)

Floris Bex

Rick Nouwen

Chris Janssen

Rosalie Iemhoff

Questions?