Artificial Intelligence

1. A Brief Introduction to AI

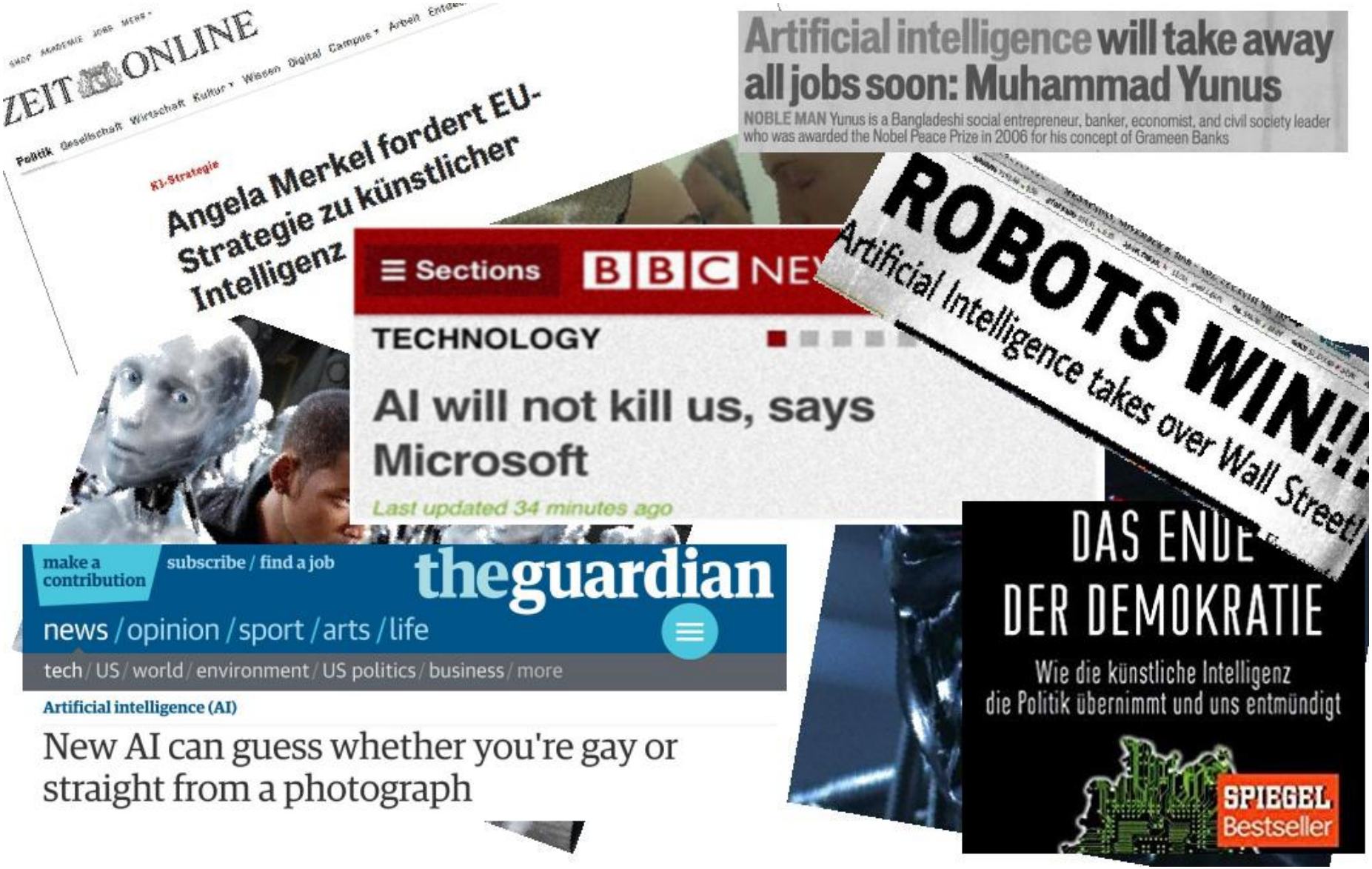
What is AI, Anyway?

Jörg Hoffmann, Daniel Fiser, Daniel Höller, Sophia Saller

SAARLAND UNIVERSITY
COMPUTER SCIENCE

Summer Term 2022

Agenda


ZEIT ONLINE
Politik Gesellschaft Wirtschaft Kultur Wissen Digital Campus Arbeit
KI-Strategie
Angela Merkel fordert EU-Strategie zu künstlicher Intelligenz
Artificial intelligence will take away all jobs soon: Muhammad Yunus
NOBLE MAN Yunus is a Bangladeshi social entrepreneur, banker, economist, and civil society leader who was awarded the Nobel Peace Prize in 2006 for his concept of Grameen Banks
Sections BBC NEWS
TECHNOLOGY
AI will not kill us, says Microsoft
Last updated 34 minutes ago
ROBOTS WIN!!!
Artificial Intelligence takes over Wall Street!
make a contribution subscribe / find a job theguardian
news / opinion / sport / arts / life
tech / US / world / environment / US politics / business / more
Artificial intelligence (AI)
New AI can guess whether you're gay or straight from a photograph
DAS ENDE DER DEMOKRATIE
Wie die künstliche Intelligenz die Politik übernimmt und uns entmündigt
SPIEGEL Bestseller

Our Agenda for This Chapter

- **AI?** What does this term even mean?
  → Spoiler: Nobody knows.
- **AI History:** How did this come about?
  → Just a little background.
- **AI Today:** What does it look like?
  → Brief research field overview.

What is Intelligence?

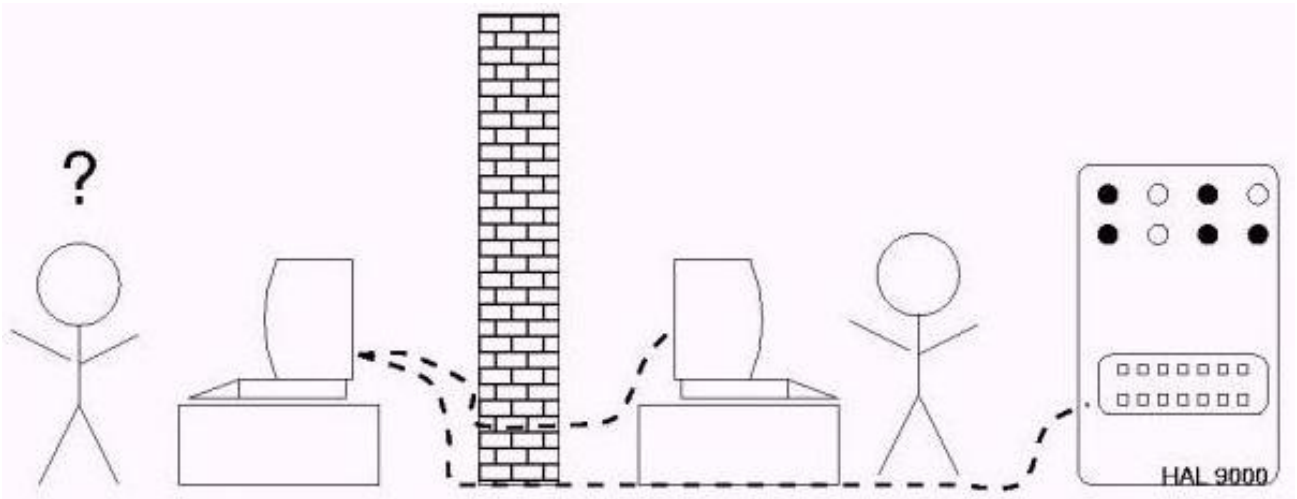
Nobody knows:

- Being good at maths? (what else?)
- Being good at Chess or Go? (what else?)
- Ability to think? (what does this mean?)
- Ability to learn? (what does this mean?)
- Creativity? (what does this mean?)
- Passing an IQ test with high marks? (go away!)

$\rightarrow$ This question has been debated in Philosophy since centuries ...

What is *Artificial* Intelligence?

Take 1: The Turing Test



- Suggested (in various forms) by Alan Turing as a measurable definition of AI.
- Yearly competitions: Loebner Prize
- Recent serial winner: Kuki (formerly Mitsuku)

What is *Artificial* Intelligence?

Take 2: Let's try to be systematic here ...

	Humanly	**Rationally**
Thinking	Cognitive Science Neural Networks? Certainly not yet!	Logics Machine Learning
Acting	Turing Test	APPLICATIONS

→ Note: Thinking is (sometimes?) a prerequisite for acting ... logics and machine learning are motivated by, and very useful in, applications!

The Four Categories: Summary

Acting Humanly: Turing Test. Not much pursued otherwise.

≈ Aeronautics: "Machines that fly so exactly like pigeons that they can even fool other pigeons".

Thinking Humanly: Cognitive Science. How do humans think, how does the human brain work.

→ Neural networks are an (extremely simple, so far) approximation.

Thinking Rationally: Logics (formalization of knowledge and deduction). Machine Learning (ML) (mathematical formulation of learning).

Acting Rationally: How to make good action choices?

→ Is what we're interested in, in practice. Encompasses logics and ML (in particular neural networks) as methods to take rational decisions.

One Recently Prominent Example of Rational Acting

AlphaGo = search + neural networks

→ We do search and ML basics here. Neural networks/deep learning are covered in the Machine Learning courses.

The History of AI

Origins: The dream of an "artificial intelligence" (broadly interpreted) is age-old (Philosophy mainly).

1956: Inception of AI at Dartmouth Workshop. John McCarthy proposes the name "Artificial Intelligence". Early enthusiasm, famous quote:

> *"It is not my aim to surprise or shock you – but the simplest way I can summarize is to say that there are now in the world machines that think, that learn and that create. Moreover, their ability to do these things is going to increase rapidly until – in the visible future – the range of problems they can handle will be coextensive with the range to which the human mind has been applied."*

60's: Early successes. "Intelligent Behavior" is shown in many demonstration systems for microworlds (Blocksworld).

70's: How to scale from microworlds to real applications?
→ Knowledge-based systems, knowledge provided by humans.

Early 80's: Commercial success of rule-based expert systems.

The History of AI, ctd.

Late 80's: Expert systems prove less promising than imagined (difficult to update/maintain, cannot learn, brittle). → "AI Winter".

90's–00's: Formalization of AI techniques and increased use of mathematics in the field. Quote from [Russell and Norvig (1995)]:

"A better understanding of the problems and their complexity properties, combined with increased mathematical sophistication, has led to workable research agendas and robust methods."

10's: Re-advent of neural networks (NN).

NN have decade-old roots. Almost forgotten in the 90's and 00's. "Sudden" success in image classification end of 00's, way better than human-coded rules. Since then, rapid successes and hype. "Deep" NN = several layers (how many? next question please).

Enablers: advanced NN architectures; lots of data; hardware.

20's: Computers rule the world? Next AI Winter?

AI Today: Sub-Areas

Modern AI is a conglomerate of highly technical sub-areas:

Search: How to effectively find solutions in problems with large search spaces (**NP**-hard and far beyond).

→ **Chapters 3–5**

CSP & SAT: General formulation and solution of search problems that involve satisfying a set of constraints.

→ **Chapters 6–7**

KR: Knowledge representation and reasoning (logic and deduction).

→ **Chapters 8–11**

Planning: General formulation and solution of search problems that involve finding goal-leading action strategies.

→ **Chapters 12 and 13** (+separate course)

AI Today: Sub-Areas, ctd.

Modern AI is a conglomerate of highly technical sub-areas:

Uncertainty: Reasoning about uncertain knowledge.

→ **Chapters 14 and 15**

ML: Machine Learning: How to learn from experience?

→ **Chapter 16** (+separate courses).

Multi-Agents: How to control/analyze systems of agents perceiving/acting individually?

→ Not considered here.

Robotics: How to control/design robots?

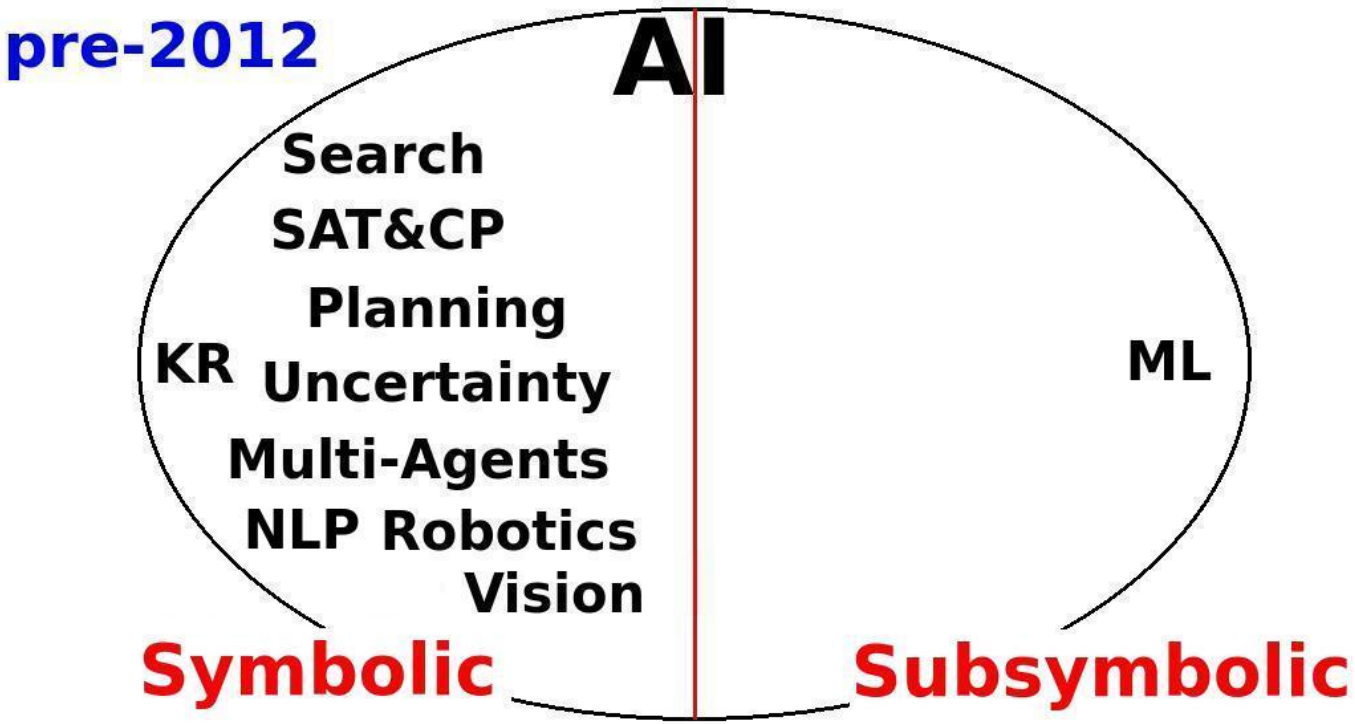
→ Not considered here.

Vision: How to interprete/analyze camera inout?

→ Not considered here (separate Courses).

→ Intimate relations to many other areas of CS. Logic Programming, Databases, Verification, Game Theory, ...

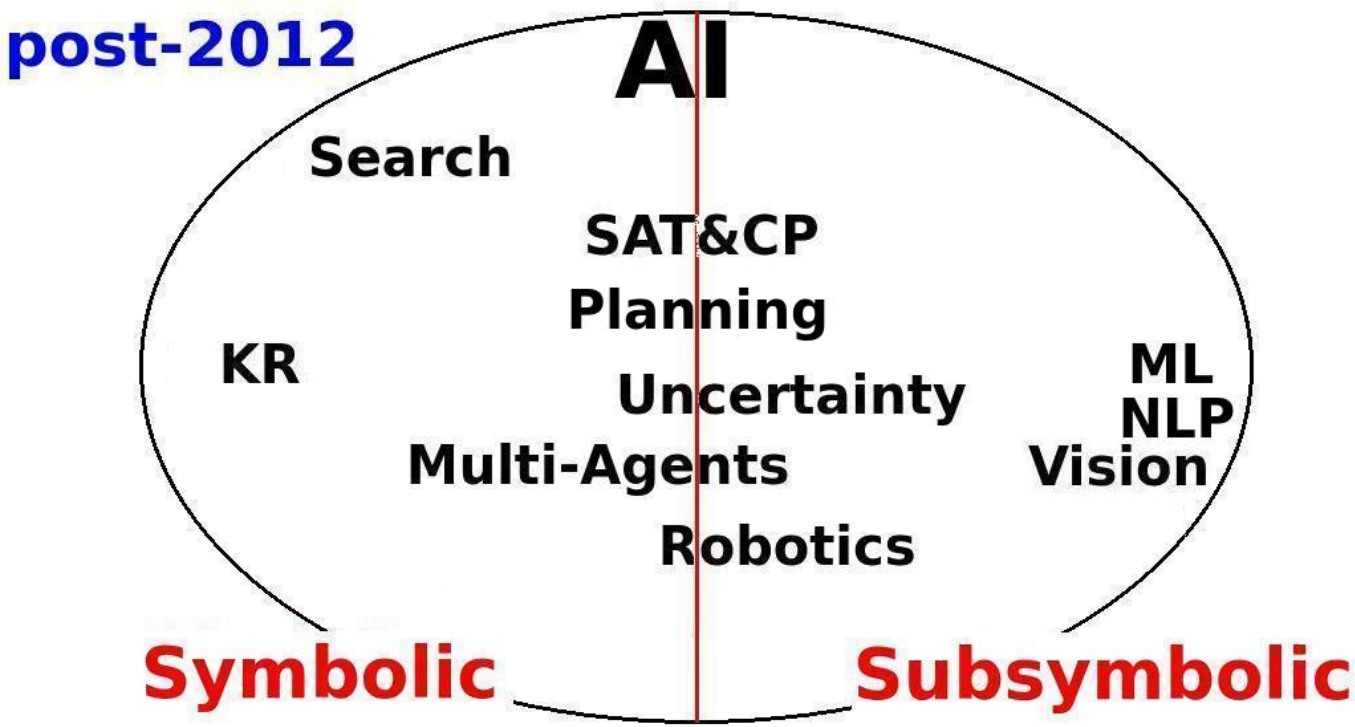
Symbolic vs. Subsymbolic AI



Symbolic: Conceptual, human-readable, formalization (model) of world behavior

Subsymbolic: Fitting of function parameters to data (world behavior observations)

Symbolic vs. Subsymbolic AI



Symbolic: Conceptual, human-readable, formalization (model) of world behavior

- Pros: instant performance, verifiability, explainability
- Cons: modeling can be costly or impossible, complexity of reasoning

Subsymbolic: Fitting of function parameters to data (world behavior observations)

- Pros: highly performant, able to tackle problems elusive for conceptual modeling
- Cons: learning curve, opaque, hyperparameters difficult to set

→ **How to combine the two?**

AI Today: A Few Applications

Go

Poker

Self-Driving Cars

Speech Recognition

Summary

- "Artificial intelligence" as an idea can be roughly classified along the dimensions thinking vs. acting and humanly vs. rationally.
- The research area of Artificial Intelligence (AI) today, as well as this course, are about "acting rationally".
- Early AI had ambitious dreams, and successes in simple problems, but then faced difficulties to scale up. Since the early 90s, AI has become more formal and systematic.
- Modern AI is a conglomerate of highly technical sub-areas, many of which have intimate relations to other areas of Computer Science.
- A key issue in AI today is symbolic vs. subsymbolic (aka data-driven) AI, and how to combine the two.

Reading

- *Chapter 1: Introduction* [Russell and Norvig (2010)].

  Content: A much more detailed account of the issues I have overviewed here.

References I

Stuart Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach*. Prentice-Hall, Englewood Cliffs, NJ, 1995.

Stuart Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach (Third Edition)*. Prentice-Hall, Englewood Cliffs, NJ, 2010.