

Artificial General Intelligence

8. Summary

Dr. Pei Wang

Dr. Patrick Hammer

Dr. Robert Johansson

Digital Futures (KTH, SU, RISE)

Stockholm, Fall 2022

Revisit the basic questions

1. **Objective:** *What* aspect of an AGI is comparable to that of a human?
2. **Possibility:** Does anyone know *whether* AGI can be achieved?
3. **Strategy:** *Which* technical approach toward AGI is mostly plausible?
4. **Morality:** *Why* should we build AGI?

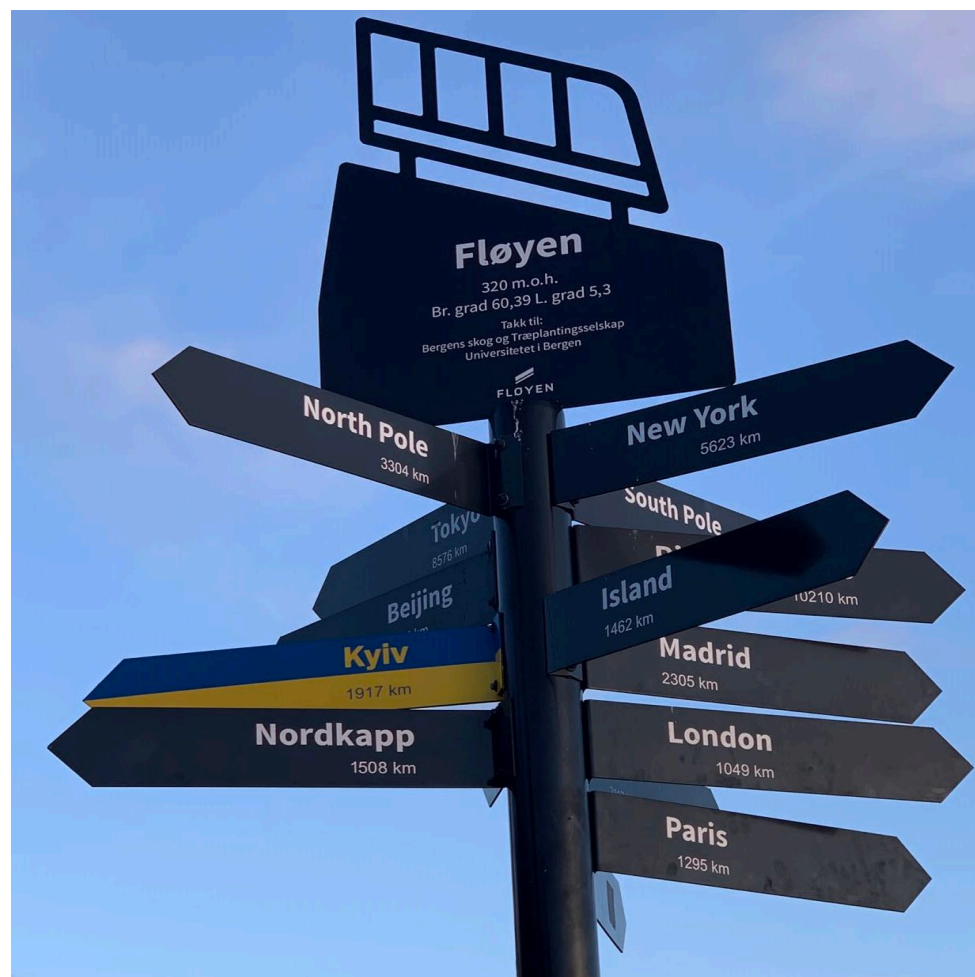
1. Objective and definition

Defining *intelligence* by abstracting *human intelligence* across boundaries such as *biological / mechanical, individual / group, etc.*

- Do not focus on human-specific details but the general principles of mind
- Do not focus on domain-specific problems but a meta-level problem
- Give the field a proper identify

1. Objective and definition (*cont.*)

This working definition of *intelligence* does not deny the values of the others, but leads the research to a different direction



2. Possibility and limitations

The practice of NARS has rebuffed claims like

- *Computer has access to syntax but not semantics, so cannot capture the meaning of concepts ...*
- *Computer always follows algorithms in problem solving, so cannot be creative or flexible ...*
- *Computer can only pursue given goals, so cannot be original or autonomous ...*
- *Computer cannot have emotion, consciousness, ...*

2. Possibility and limitations (cont.)

NARS will not meet expectations of AGI like

- *AGI means Singularity and Superintelligence that are beyond human comprehension ...*
- *AGI will solve all problems that can be solved by human, so will take over all jobs from us ...*
- *AGI will surpass special-purpose techniques ...*

NARS will not be a commercial product, though parts of it (or a customized version) can be

3. Strategy and technique

- NARS has been developed according to its objective through incremental designing and implementing
- This theory-model-system gets ideas from many theories and techniques without being a hybrid of them
- Other models and techniques can be used as “tools” or “subsystems” for specific purposes

3. Strategy and technique (*cont.*)

- The theory starts by establishing a new standard of *rationality* in *realistic* situations
- The model treats *meaning* and *truth-value* according to experience-grounded semantics
- The logic extends traditional models beyond *binary deduction* among *atomic* terms
- The control accomplishes *case-by-case* problem-solving through dynamic resource allocation

4. Morality and safety

- Beside the common issues in ethics of technology, adaptive systems have the special issue of *plasticity* or *nurture*
- The intelligence and morality of a system are largely independent of each other
- The moral and ethical problems of NARS will mostly be handled by a collaboration of *design*, *configuration*, *education*, *socialization*, *evolution*, ...

4. Morality and safety (cont.)

If the safety and morality of AGI cannot be guaranteed, why to take this risk?

- This is a cross-disciplinary frontier in the study of *thinking, cognition, mind, mentality, ...*
- This research has the potential of inventing revolutionary technologies that solve vital problems of human society

It is a matter of weighing the pros and cons when making a decision, as intelligence requires ...

Status and future

- The field of AGI is still in the stage where multiple objectives and approaches compete with each other
- NARS progresses in an increasing speed in the recent years, though there are still many problems to be solved
- Various forms of involvement and cooperation are welcome

Suggested Readings

- Pei Wang, [The basic questions about AGI](#), *Digital Futures* Lecture, Sept. 8, 2022
 - Pei Wang, Non-Axiomatic Logic, Chapter 14
 - Pei Wang, Rigid Flexibility, Chapter 13-14
-

[Personal page] <https://cis.temple.edu/~pwang/>

[Temple team] <https://cis.temple.edu/tagit/>

[Stockholm lab] <https://stockholm-agi-lab.github.io/>