Theoretical framework for description and modeling of heterogeneous cognitive systems

Julian Zubek, Łukasz Jonak, Dariusz Plewczyński, Joanna Rączaszek-Leonardi

HILL meeting 07/02/2020

Motivation

Our understanding of the world is not “ours” alone! We live as part of heterogeneous cognitive systems.

human - AI, AI - AI, human - human

- With information/communication technologies the systems scale-up.
- … but our intuitions do not.

Need for a renaissance of systemic thinking.

Not just single agents but agents functioning in highly structured environment…

Systemic framework

- Do not rely on a general theory of cognition, intelligence, etc.
- Do not focus on internal structure and mechanisms of individual agents.
- Look for systemic organization: how agents coordinate to act as a whole.
- Be pragmatic: think of the modeler’s needs.

Not a unified theory, but a general framework: main concepts may be differently operationalized in different domains.

Agent and degrees of freedom

Agent:

- Entity acting in a system.
- Characterised by a minimal definition of “operative intentionality” (Merleau-Ponty), the set of agent’s choices of its interactions with environment.

Agency and Degrees of Freedom:

- Agent’s agency – organizations agent’s choices, degrees of freedom.
- DoF: independent parameters required to describe dynamics of a system. eg.:
  - Position and velocity in physical systems;
  - The ways model can ‘pivot around’ data;
  - The number of symbols in vocabulary;
  - Number of courses of action that can be taken by agents.

System boundary

Decision what to include and what to exclude from a system – made by the modeler (depends on the particular needs).

But some choices are more natural:

- Machines operate under controlled conditions as semi-closed systems and are often shielded from the environment.
- Many natural network display clustering properties.
- Living organisms are open systems, which actively construct their boundaries (operative closure in autopoiesis theory).
- Social systems may also be autopoietic in some sense.

System boundary is a natural point of reference when describing system’s activity.

Actions and constraints on degrees of freedom

Actions are organization's of actor's degrees of freedom.

Activity of an agent (related to the boundary of the system):

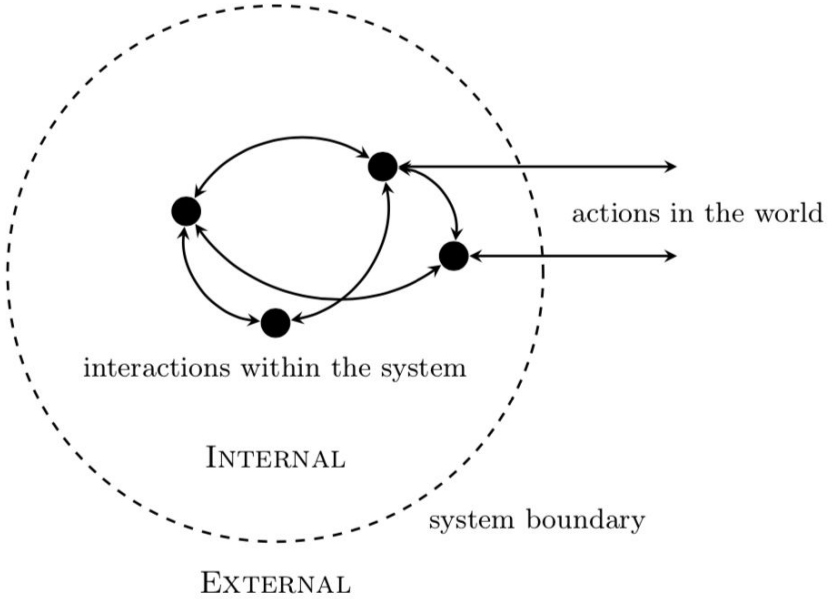
- Internal, affects the system itself
- External, affects outside world

Constraints are circumstances restricting/affording degrees of freedom:

- Environment
- Other agents
- Individual, cultural, biological history
- *Can be shared by agents*

- Internal: regulate internal activities
- External: regulate external activities

Anatomy of a system



Agents acting within a system. Agents are represented as dots, their activity is divided into internal and external based on the system boundary.

Simple example: shepherd and shepherd's dog

Agents: shepherd and his dog.

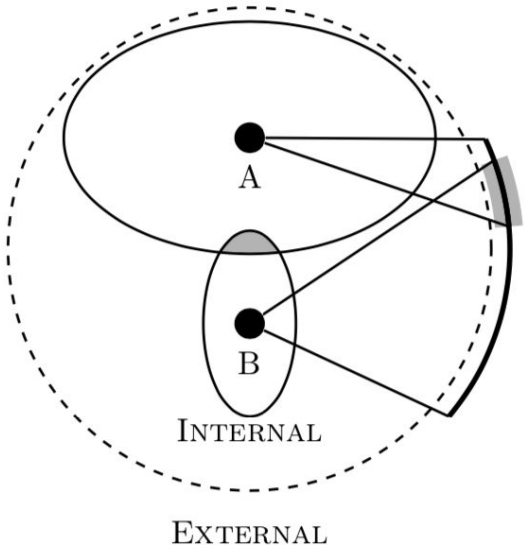
Environment: sheep, terrain, weather conditions, …

Shepherd giving commands to the dog: internal activity.

Dog herding sheep: external activity.

Dynamics of agents' movement (e.g., their distance from each other) regulated by constraints.

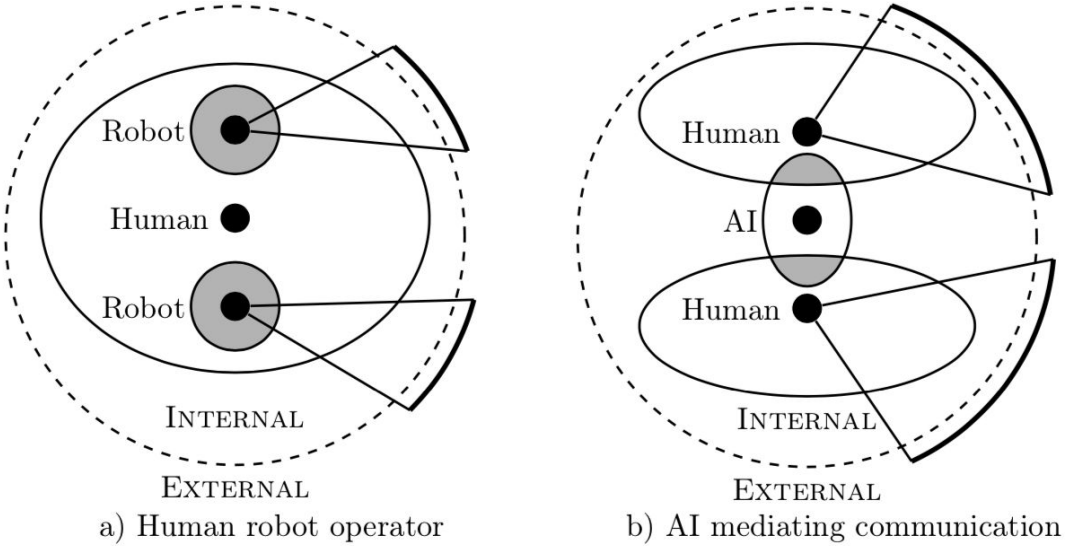
Assembling agents into systems



Specialized agents with different scopes of activity. **Ellipses**: scopes of *internal* activity, **arcs**: scopes of *external* activity. **Shaded area**: the amount of *shared constraints*. Agent A is active mostly externally, agent B is active mostly internally.

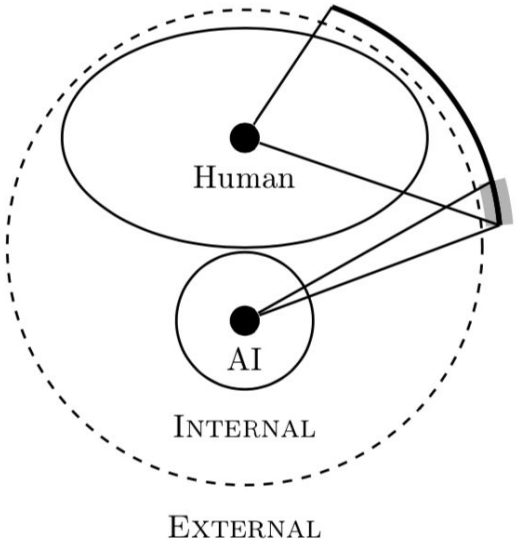
- Whenever two agents have different sets of **external** constraints there is room for synergy between them
- Agents with overlapping sets of constraints can perform the same **external** activities and act as a replacement for each other. Introducing redundancy into the system may increase its robustness to failure of individual agents.
- **Internal** constraints have to be shared/compatible for coordination of actions, allowing for “pooling” their resources in the form of respective ranges of degrees of freedom
- Agents with non-overlapping set of **internal** constraints are usually unable to collaborate effectively

Various compositions of internal activities in heterogeneous systems



Two scenarios of mediating interaction: a) human controlling and coordinating actions of artificial agents, thus effectively acting as an intermediary between them, b) artificial agent filtering and moderating communication between humans, helping them focus on relevant dimensions.

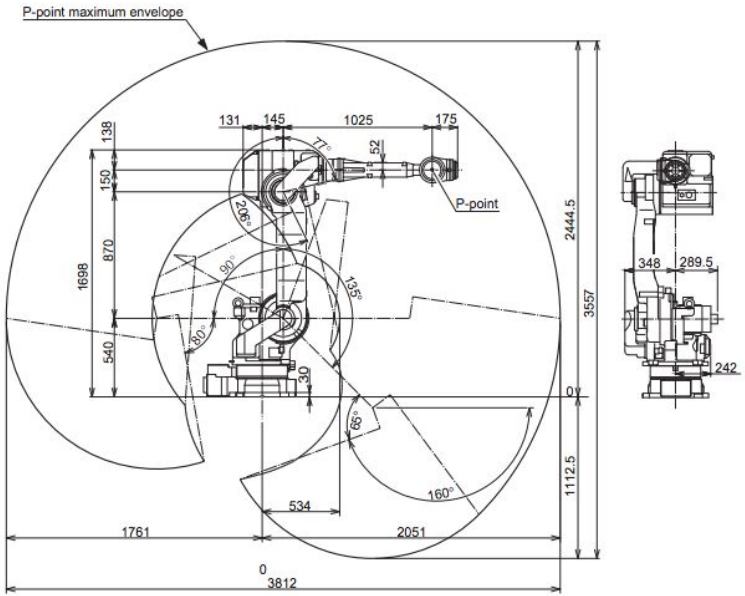
Human and Artificial agents



Human intelligence usually has more degrees of freedom and is less constrained than AI.

- Artificial and natural systems differ in ranges of their variability and the composition of the internal and external constraints as well as the capacity of their organization in the face of interaction with the environment.
- Different number of internal degrees of freedom of the agents, different characteristics of internal constraints (very specific and limiting versus broad, flexible), and different amount of overlap in the external constraints (separated contexts of computer programs versus situational context shared by humans).

Humans and AI - Enveloping



Working envelope of industrial robot

Enveloping: the need to modify the environment so that it is compatible with artificial actor's limited degrees of freedom.

- Human-made environment—digital (the Internet), social or built (for example, cities), is transformed to better accommodate the needs of more restricted artificial agents (L. Floridi).
- The job of a human becomes to mediate between the enveloped AI world and everyday reality.
- With the proliferation of AI systems, there is a tendency to impose strong constraints on a collective system to reduce it to its least common denominator—to more limited DoF repertoire of an artificial agent.

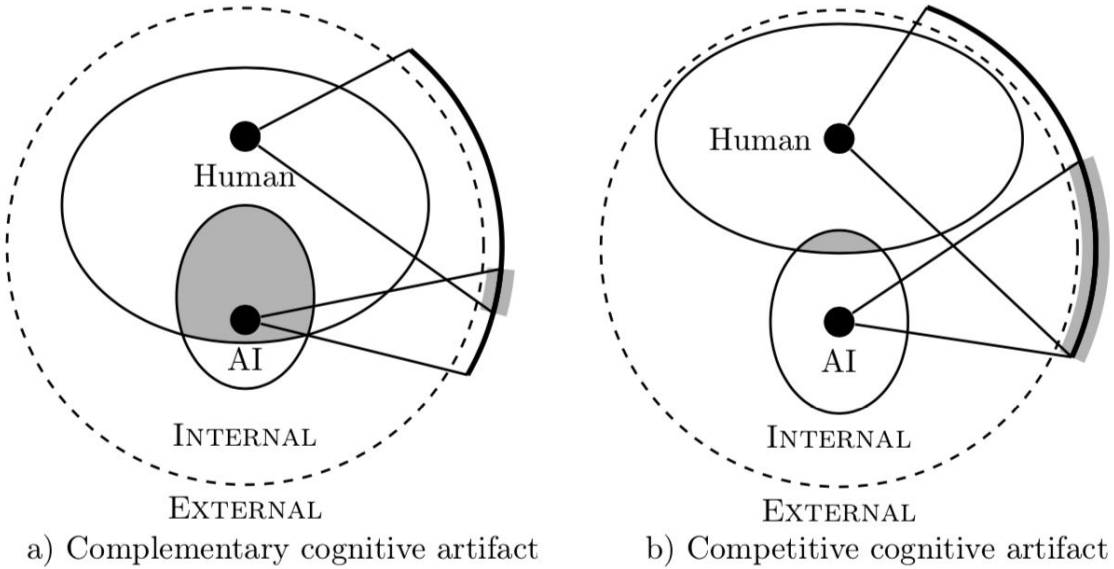
Humans and AI: Complementary vs Competitive Cognitive Artifacts

Complementary vs Competitive cognitive artifacts (CAs) (D.Krakauer)

- Complementary CA: transparent to user, enhances system's performance in some domain; when CA removed, user still can perform the task (sometimes with enhanced performance, as with the abacus).

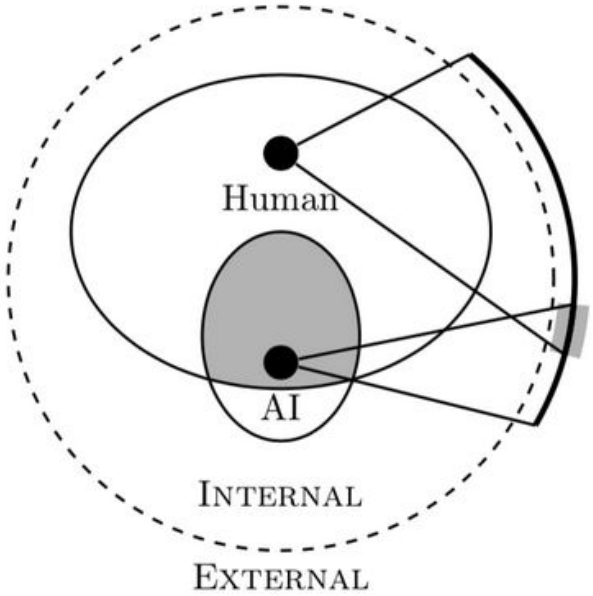
- Competitive CA: opaque to user, enhances system's performance in some domain, when CA removed, user's performance hindered.

Complementary vs Competitive CA's, cont.



AI agents as **complementary** (small overlap in the set of external constraints, large overlap in the set of internal constraints) and **competitive** (small overlap in the set of internal constraints, large overlap in the set of external constraints) cognitive artifacts.

How human-AI hybrid systems should be assembled?



AI remodelling human relationship with the environment needs to be counterbalanced with binding (as much as possible) AI's and human internal degrees of freedom.

It is about control, comprehension, participation in/of the system.

- Ethical AI
- Explainable AI
- Interfaces (NeuraLink)


PEOPLE'S DAILY
Using their left hands and a mental image
of abacus calculation, these young kids can
do math work at a tremendous speed.