Eliezer Yudkowsky - Levels of Organization in General Intelligence (2007)

Tom Rochette <tom.rochette@coreteks.org>

November 2, 2024 — 36c8eb68

0.1 Context

0.2 Learned in this study

- Concept creation is limited by our capacity to create feature kernels (we can cluster "red", but we can't cluster hue or saturation, or can we with enough training?)

0.3 Things to explore

- What is the hierarchy of components that composes an AGI?

1 Overview

- The most important skill an AI develop can have is knowing what not to program

1.1 Mistakes in AI

- Implementing a process too close to the token level - trying to implement a high-level process without implementing the underlying layers of organization
- The failure to distinguish between things that can be hardwired and things that must be learned

2 Notes

2.1 1 Foundations of General Intelligence

- What is intelligence? A brain with a hundred billion neurons and a hundred trillion synapses; a brain in which the cerebral cortex alone is organized into 52 cytoarchitecturally distinct areas per hemisphere
- Intelligence is the complex expression of a complex set of principles
- Intelligence is a supersystem composed of many mutually interdependent subsystems - subsystems specialized not only for particular environmental skills but for particular internal functions
- Intelligence requires more than passive correspondence between internal representations and sensory data, or between sensory data and reality
- Intelligence in the fully human sense requires the ability to manipulate the world by reasoning backward from a mental image of the desired outcome to create a mental image of the necessary actions
- A new (genetic) adaptation requiring the presence of a previous adaptation cannot spread unless the prerequisite adaptation is present in the genetic environment with sufficient statistical regularity to make the new adaptation a recurring evolutionary advantage
- Evolution often finds good ways, but rarely the best ways
- Evolution is a useful inspiration but a dangerous template

- Because evolution lacks foresight, complex functions cannot evolve unless their prerequisites are evolutionary advantages for other reasons
- A deliberately designed AI is likely to begin as a set of subsystems in a relatively primitive and undeveloped state, but nonetheless already designed to form a functioning supersystem
- Where the human line developed from very complex non-general intelligence into very complex general intelligence, a successful AI project is more likely to develop from a primitive general intelligence into a complex general intelligence
- The abstract knowledge that biological neurons implement human intelligence does not explain human intelligence
- The most common paradigms of traditional AI - search trees, neural networks, genetic algorithms, evolutionary computation, semantic nets - have in common the property that they can be implemented without requiring a store of preexisting complexity
- Today we have at least one reason to believe that nonsensory intelligence is a bad approach; we tried it and it didn't work
- Yudkowsky argues strongly for "supersystems," but he does not believe that "supersystems" are the necessary and sufficient Key to AI
- General intelligence requires the right supersystem, with the right cognitive subsystems, doing the right things in the right way
- Humans are not intelligent by virtue of being "supersystems," but by virtue of being a particular supersystem which implements human intelligence

2.2   2 Levels of Organization in Deliberative General Intelligence

2.3   2.1 Concepts: An Illustration of Principles

- DGI hypothesizes that visualization occurs through a flow from high-level feature controllers to low-level feature controllers, creating an articulated mental image within a sensory modality trhough a multistage process that allows the detection of conflicts at higher levels before proceeding to lower levels
- The final mental imagery is introspectively visible, but the process that creates it is mostly opaque
- While the cat has roughly 10^6 fibers from the lateral geniculate nucleus to the visual cortex, there are approximately 10^7 fibers running in the opposite direction
- No explanatory consensus currently exists for the existence of the massive corticothalamic feedback projects, though there are many competing theory
- In concept combination, a few flashes of the intermediate stages of processing may be visible as introspective glimpses - especially those conflicts that arise to the level of conscious attention before being resolved automatically
- Almost all of the internal complexity of concepts is hidden away from human introspection, and many theories of AI (even in the modern era) thus attempt to implement concepts on the token level, e.g., "lightbulb" as a raw LISP atom
- DGI belongs to the existing tradition that ask, not "How do we ground our semantic nets?", but rather "What is the underlying stuff making up these rich high-level objects we call 'symbols'?"
- The "frame problem" - sometimes also considered a form of the "commonsense problem" - in which intelligent reasoning appears to require knowledge of an infinite number of special cases
- Consider a CPU which adds two 32-bit numbers
- An "expert system" that encodes a vast semantic network containing the "knowledge" that 2+2=4, 21+16=37 and so on would require a lookup table of 18 billion billion entries
- Symbol grounding: Linking G0025 to hamburger or Johnny Carson
- Yudkowsky believes that many (though not all) cases of the "commonsense problem" or "frame problem" arise from trying to store all possible descriptions of high-level behaviors that, in the human mind, are modeled by visualizing the lower level of organization from which those behaviors emerge
- The visual cortex "knows" about edge detection, shading, textures of curved surfaces, binocular disparities, color constancy under natural lighting, motion relative to the plane of fixation, and so on
- Even if an AI needs programmer-created concepts to bootstrap further concept formation, bootstrap concepts should be created using programmer-directed tool versions of the corresponding AI subsystems,

and the bootstrap concepts should be replaced with the AI-formed concepts as early as possible
- Two potential problems emerging from the use of programmer-created content are opacity and isolation
    - Opacity refers to the potential inability of an AI's subsystems to modify content that orginated outside the AI
        * If a programmer is creating cognitive content, it should at least be the kind of content that the AI could have created on its own; it should be content in a form that the AI's cognitive subsystems can manipulate
    - Isolation means that if a concept, or a piece of knowledge, is handed to the AI on a silver platter, the AI may be isolated from the things that the AI would have needed to learn first in order to acquire that knowledge naturally, in the course of building up successive layers of understanding to handle problems of increasing complexity

2.4   2.2 Levels of Organization in Deliberation

- 5 distinct layers of organization
    1. Source code and data structures (human equivalent is neurons and neural circuitry)
    2. Sensory modalities
    3. Concepts (aka categories or symbols)
    4. Thoughts
    5. Deliberation (explanation, prediction, planning, design, discovery and other activities used to solve knowledge problems in the pursuit of real-world goals)

2.5   2.3 The Code Level

- Composed of functions, classes, modules, packages; data types, datastructures, data repositories
- Trying to create general AI means solving the programming problem of creating a mind that solves problems
- Good programming tools help with AI development but do not help with AI
- Writing a great deal of exploratory code means that the IDEs with refactoring support and version control are important, and that modular code is even more important than it is usually
- Human cognition may be massively parallel on the lower levels, but the overall flow of cognition is still serial
- We know it is possible to evolve a general intelligence that runs on a hundred trillion synapses with characteristic limiting speeds of approximately 200 spikes per second
- It seems likely that the brain uses a massively parallel algorithm at one point or another of its operation; memory association is simply a plausible candidate
- The human brain is massively parallel because massive parallelism is the only way to do anything in 200 clock ticks
- Instead of trying to reimplement the cognitive process of association as it developed in humans, we might instead ask: What would this cognitive subsystem look like, if it had evolved on hardware instead of wetware?
- Smooth fitness landscapes are a useful, necessary, and fundamental outcome of evolution
- Smooth fitness landscapes imply, among other things, that a small perturbation in the program code (genetic noise), in the input (environmental noise), or in the state of the executing program (processor noise), is likely to produce at most a small degradation in output quality
- In most human-written code, a small perturbation of any kind usually causes a crash
- Human-written programs start out as high-level goals which are translated, by an extended serial thought process, into code
- There are important features that are also supported by biological neural network - that are "natural" to neural substrate. These features probably include:
    - Optimizing for recurring problems
    - Completion of partial patterns
    - Similarity recognition (detection of static pattern repetition)
    - Recurrence recognition (detection of temporal repetition)

- – Clustering detection, cluster identification, and sorting into identified clusters
- – Training for pattern recognition and pattern completion
- – Massive parallelism
- Contemporary hardware has better support for:
  - – Reflectivity and execution traces
  - – Lossless serialization (storage and retrieval) and lossless pattern transformations
  - – Very-high-precision quantitative calculations
  - – Low-level algorithms which involve extended iteration, deep recursion, and complex branching
  - – Massive serialism, the ability to execute hundreds of millions of sequential steps per second
- Programming is not all the work of AI, perhaps not even most of the work of AI; much of the effort needed to construct an intelligence will go into prodding the AI into forming certain concepts, undergoing certain experiences, discovering certain beliefs, and learning various high-level skills
- The most important skill an AI develop can have is knowing what not to program

2.6 2.4 The Modality Level

2.6.1 The Evolutionary Design of Modalities in Humans

- Visual preprocessing begins in the first layer of the retina, which converts the raw intensities into center-surround gradients, a representation that forms the basis of all further visual processing
- After several layers of retinal processing, the final retinal layer is composed of a wide variety of ganglion types that include directionally selective motion detectors, slow-moving edge detectors, fast movement detectors, uniformity detectors, and substractive color channels
- Information streams to area V1, the primary visual cortex, which begins featuer extraction for information about motion, orientation, color and depth
- From primary visual cortex the information stream continues, making its way to the higher visual areas, V2 through V6
- Beyond the visual cortex, the information stream continues to temporal areas (object recognition) and parietal areas (spatial processing)
- The general lesson learned from the human visual modality is that modalities are not microtheories, that modalities are not flat representations of the pixel level, and that modalities are functionally characterized by successive layers of successively more elaborate feature structure
- In modalities, each additional layer of feature detectors makes use of the information provided by the first layer of feature detectors
- Evolution is a good teacher but a poor role model
- Evolution, in constructing a system incrementally, tends to construct linear sequences or ascending layers of smooth operations

2.6.2 The Human Design of Modalities in AI

- Should humanlike 3D vision be one of the first modalities attempted?
- For an AI's initial learning experiences, Yudkowsky would advocate placing the AI in complex virtual environments, where the virtual environments are internal to the computer but external to the AI
- Shifting to microworlds does require that experience in the microworlds reproduce functionally relevant aspects of experience in real life, including unpredictability, uncertainty, real-time process control, holonic (part-whole) organization, et cetera
- Three possible modalities that come to mind as reasonable for a very primitive and early-stage AI, in ascending order of implementational difficulty, would be:
  1. A modality for Newtonian billiard balls
  2. A modality for a 100x100 "Go" board
  3. A modality for some type of interpreted code (a metaphorical "codic cortex")
- Yudkowsky recommends eliminating environmental reconstruction as a complexity source in early AI development
- Sensory modalities are about invariants and not just regularities
- When considering which features to extract, the question I would ask is not "What regularities are

found in code?" but rather "What feature structure is needed for the AI to perceive two identical algorithms with slightly different implementations as 'the same piece of code'?" Or more concretely: "What features does this modality need to extract to perceive the recursive algorithm for the Fibonacci sequence and the iterative algorithm for the Fibonacci sequence as 'the same piece of code'?"

- The modality level is about invariants rather than regularities and identities rather than categories

2.7   2.5 The Concept Level

- Concept refers to the mental stuffs underlying the words that we combine into sentences; concepts are the combinatorial building blocks of thoughts and mental imagery
- A "concept kernel" is the pseudo-sensory pattern produced by abstracting from sensory experience
- A programmer seeking a good representation for concept kernels must find a representation that simultaneously fulfills these requirements:
    - The kernel representation can be satisfied by and imposed on referents in a sensory modality
    - The kernel representation or concept representation contains the internal structure needed for faceted concept combination
    - It is computationally tractable to abstract new kernel representations using sensory experience as raw material

2.7.1   The Substance of Concepts

- One of the meanings of "abstraction" is "removal"
- To create a concept is to generalize
- To generalize is to lose information
- Yudkowsky suspect that concepts do not have independent procedures for satisfaction and imposition; Yudkowsky also suspects that neither satisfaction nor imposition are the product of reinforcement learning on a fully general procedure. Rather, Yudkowsky suspects that a concept kernel consists of a pattern in a representational related to (but not identical with) the representation of sensory imagery, that this pattern is produced by transforming the experiences from which the concept is abstracted, and that this pattern interacts with the modality to implement both concept satisfaction and concept imposition
- Even though an underlying "rule" is computationally very simple, it may be unlikely that a human will create a concept that directly incorporates the rule; it may even be impossible for a human to abstract a kernel that performs this simple computation (think of a a set of all colors with the same hue or saturation)

2.7.2   Stages in Concept Processes

- Concept formation is a multi-stage process
- After the creation of the concept and concept kernel(s), it would then be possible for the AI to notice concept-concept relations, such as supercategory and subcategory relations
- After a concept has been formed, the new concept must be integrated into the system
- Yudkowsky conjectures that matching imagery against large concept sets will be one of the most computationally intensive subprocesses in AI, perhaps the most expensive subprocess

2.7.3   Complex Concepts and the Structure of "Five"

- The most important reason for decomposability is that concepts with organized internal structures are more mutable
- A concept with internal structure or procedural structure, created by the AI's own thought processes in response to experience, is mutable by the AI's thought processes in response to further experience

2.7.4   Experiential Pathways to Complex Concepts

- The full path to fiveness would probably involve:

- Learning physical continuity
- Learning unique correspondence
- Learning complete mapping
- Complete mapping using the unique-correspondence relation, also known as one-to-one mapping
- With one-to-one mapping, it is possible for the AI to notice that all the answers on a challenge task are related to a common prototype
- A deliberative intelligence must build up complex concepts from simple concepts, in the same way that evolution builds high-level feature detectors above low-level feature detectors, or builds organs using tissues, or builds thoughts over concepts or modalities

2.7.5 Microtasks

- Forming a complex concept requires an incremental path to that complex concept - a series of building-block concepts and precursor concepts so that the final step is a leap of manageable size
- Under the microtask developmental model, this would be implemented by a series of microtasks of ascending difficulty and complexity, in order to coax the AI into forming the precursor concepts leading up to the formation of complex concepts and abstract concepts

2.7.6 Interactions on the Concept Level

- To summarize some of the complexity bound up in concept-concept interactions:
  - Concepts are associated with other concepts. Activating a concept can "prime" a nearby concept, where "priming" is usually experimentally measured in terms of decreased reaction time
  - Nearby concepts may sometimes "slip" under cognitive pressures
  - Concepts, in their role as categories, share territory
  - Concepts, in their role as categories, have supercategory and subcategory relationships
  - Concepts possess transformation relations (a -> b, 1 -> 2)
  - Concepts act as verbs, adjectives, and adverbs as well as nouns
  - Concepts, in their role as symbols with auditory, visual, or gestural tags, play a fundamental role in both human communication and internal human conceptualization

2.8 2.6 The Thought Level

- Thoughts are structures of combinatorial concepts that alter imagery within the workspace of sensory modalities
- Modalities are wired; concepts are learned; thoughts are invented

2.8.1 Mental Imagery

- Abstract reasoning is a means by which inductively acquired generalizations can be used in deductive reasoning
- Confidence is a percept that sums up a network of reliances on uncertainty sources
- An AI might best begin with separate percepts for "humanlike" confidence and "arithmetical" confidence

2.8.2 The Origin of Thoughts

- Thoughts are the cognitive events that change mental imagery
- Where do thoughts come from?
  - First, the production of suggested thoughts
  - Second, the selection of thoughts that appear "useful" or "possibly useful" or "important" or otherwise interesting

2.8.3 Beliefs

- Beliefs are learned, specific, stored, and recurrent

- Beliefs are acquired through two major sources, induction and deduction, respectively referring to generalization over experience, and reasoning from previous beliefs
- The strongest beliefs hve both inductive and deductive support: deductive conclusions with experiential confirmation, or inductive generalizations with causal explanations
- Concepts are either useful or not useful, but beliefs are either true or false

2.8.4 Coevolution of Thoughts and Language: Origins of the Internal Narrative

- The key difficulty in language - the step that is awkward for other species - is not the ability to associate signals; primates (and rats, for that matter) can readily associate a perceptual signal with a required action or a state of the world. The awkward step is for a signal to evoke a category as abstract imagery, apart from immediate sensory referents, which can bind to a referent later encountered
- A de novo mind design could disentangle intelligence from language

2.9 2.7 The Deliberation Level

2.9.1 From Thoughts to Deliberation

- In humans, higher levels of organization are generally more accessible to introspection
- In humans concepts are smoothly flexible and expressive because they arise from modalities; thoughts are smoothly flexible and expressive because they arise from concepts

2.9.2 The Dimensions of Intelligence

- Yudkowski would say beliefs are useful because they are true, not "true" because they are useful
- But usefulness is a stronger and more reliable test of truth; it is harder to cheat
- 4 successively stronger grades of binding between a model and reality:
  - A sensory binding: when there is a mapping between cognitive content in the model and characteristics of external reality
  - A predictive binding: when a model can be used to correctly predict future events
  - A decisive binding: when the model can predict the effects of several possible actions on reality, and choose whichever action yields the best result under some goal system
  - A manipulative binding: when the AI can describe a desirable future with subjunctive imagery, and invent a sequence of actions which leads to that future
- 3 successive grades of variable complexity:
  - A discrete variable: selected from a bounded set which is computationally small
  - A quantitative variable: selected from the set of real numbers, or from a computationally large set which approximates a smoothly varying scalar quantity
  - A patterned variable: composed of a finite number of quantitative or discrete elements

2.9.3 Goals

- Individual organisms are best regarded as adaptation-executers rather than fitness-maximizers

2.9.4 Activities of Intelligence: Explanation, Prediction, Discovery, Design

- Explanation fills holes in knowledge about the past
- Prediction fills holes in knowledge about the future
- Discovery fills holes in knowledge about the present
- Design fills gaps in the mental model of a tool
- Planning fills gaps in a model of future strategies and actions
- Explanation, prediction, discovery, and design may be employed in the pursuit of a specific real-world goal, or as an independent pursuit in the anticipation of the resulting knowledge being useful in future goals - "curiosity"
- What all the activities have in common is that they involve reasoning about a complex, holonic model of causes and effects

- This gives us the three elements of DGI:
  - The what of intelligence
  - The why of intelligence
  - The how of intelligence

2.9.5 General Intelligence

- Not only is it impossible to duplicate general intelligence through the sum of (specialized) subsystems, Yudkowsky suspect it is impossible to achieve humanlike performance in most single contexts using specialized AI

2.9.6 Self

- The degrees of binding between a model and reality are sensory, predictive, decisive, and manipulative
- An intelligent mind navigates a universe with four major classes of variables:
  - random factors
  - variables with hidden values
  - the actions of other agents
  - the actions of the self

2.10 3 Seed AI

- There is a fundamental rift between evolutionary design and deliberative design
- Mutations are atomic; recombinations are random; changes are made on the genotype's lowest level of organization (flipping genetic bits); the grain size of the component tested is the whole organism; and the goodness metric operates solely through induction on historically encountered cases, without deductive reasoning about which contextual factors may later change
- Evolution simply cannot handle simultaneous dependencies, unless individual changes can be deployed incrementally, or multiple phenotypical effects occur as the consequence of a single genetic change
- For humans, planning coordinated changes is routine; for evolution, impossible
- Until a stably functioning cognitive supersystem is achieved, only the nondeliberative intelligence exhibited by pieces of the system will be available
- When the AI is "young" it is likely to be limited to understanding simple elements of a system, or small organizations of elements, and only where clear-cut goal contexts exist (probably programmer-explained)
- A seed AI is an AI designed for self-understanding, self-modification, and recursive self-improvement
- The consequences of seed AI (such as true recursive self-improvement) only show up after the AI has achieved significant holonic understanding and general intelligence

2.11 3.1 Advantages of Minds-in-General

- Humans definitely possess the following advantages, relative to current AIs:
  - We are smart, flexible, generally intelligent organisms with an enormous base of evolved complexity, years of real-world experience, and 10^{14} parallelized synapses, and current AIs are not
- Humans probably possess the following advantages, relative to intelligences developed by humans on forseeable extensions of current hardware:
  - Considering each synaptic signal as roughly equivalent to a floating-point operation, the raw computational power of a human is enormously in excess of any current supercomputer or clustered computing system
  - Human neural hardware - the wetware layer - offers built-in support for operations such as pattern recognition, pattern completion, optimization for recurring problems, etc.; this support was added from below, taking advantage of microbiological features of neurons, and could be enormously expensive to simulate computationally to the same degree of ubiquity
  - With respect to the holonically simpler levels of the system, the total amount of "design pressure" exerted by evolution over time is probably considerably in excess of the design pressure that a resonably-sized programming team could expect to personally exert

- – Humans have an extended history as intelligences; we are proven software
- Current computer programs definitely possess these mutually synergetic advantages relative to humans:
  - – Computer programs can perform highly repetitive tasks without boredom
  - – Computer programs can execute complex extended tasks without making that class of human errors caused by distraction or short-term memory overflow in abstract deliberation
  - – Computer hardware can perform extended sequences of simple steps at much greater serial speeds than human abstract deliberation or even human 200 Hz neurons
  - – Computer programs are fully configurable by the general intelligences called humans (Evolution, the designer of humans, cannot invoke general intelligence)

2.12   3.2 Recursive Self-enhancement

- If computational subsystems give rise to cognitive talents, and cognitive talents plus acquired expertise give rise to domain competencies, then self-improvement is a means by which domain competencies can wrap around and improve computational subsystems
- Recursive self-improvement is a consequence of seed AI, not a cheap way to achieve AI
- Epochs for holonic understanding and holonic programming:
  - – First epoch: The AI can transform code in ways that do not affect the algorithm implemented ("Understanding" on the order of an optimizing compiler; i.e. not "understanding" in any real sense)
  - – Second epoch: The AI can transform algorithms in ways that fit simple abstract beliefs about the design purposes of code. That is, the Ai would understand what a stack implemented as a linked list and a stack implemented as an array have in common
  - – Third epoch: The AI can draw a holonic line from simple internal metrics of cognitive usefulness (how fast a concept is cued, the usefulness of the concept returned) to specific algorithms. Consequently the AI would have the theoretical capability to invent and test new algorithms. This does not mean the AI would have the ability to invent good algorithms or better algorithms, just that invention in this domain would be theoretically possible
  - – Fourth epoch: The AI has a concept of "intelligence" as the final product of a continuous holonic supersystem. The AI can draw a continuous line from its abstract understanding of intelligence to its introspective understanding of cognition to its understanding of source code and stored data. The AI would be able to invent an algorithm or cognitive process that contributes to the intelligence in a novel way and integrate that process into the system
- Epochs for sparse, continuous, and recursive self-improvement:
  - – First epoch: The AI has a limited set of rigid routines which it applies uniformly. Once all visible opportunities are exhausted, the routines are used up. This essentially analogous to the externally driven improvement of an optimizing compiler. An optimizing compiler may make a large number of improvements, but they are not self-improvements, and they are not design improvements. An optimizing compiler tweaks assembly language but leaves the program constant
  - – Second epoch: The cognitive processes which create improvements have characteristic complexity on the order of a classical search tree, rather than on the order of an optimizing compiler. Sufficient investments of computing power can sometimes yield extra improvements, but it is essentially an exponential investement for a linear improvement, and no matter how much computing power is invested, the total kind of improvements conceivable are limited
  - – Third epoch: Cognitive complexity in the AI's domain competency for programming is high enough that at any given point there is a large number of visible possibilities for complex improvements, albeit perhaps minor improvements. The AI usually does not exhaust all visible opportunities before the programmers improve the AI enough to make new improvements visible. However, it is only programmer-driven improvements in intelligence which are powerful enough to open new volumes of the design space.
  - – Fourth epoch: Self-improvements sometimes result in genuine improvements to "smartness,", "creativity," or "holonic understanding," enough to open up a new volume of the design space and make new possible improvements visible
- Epochs for relative human-driven and AI-driven improvement:

- First epoch: The AI can make optimizations at most on the order of an optimizing compiler, and cannot make design improvements or increase functional complexity. The combination of AI and programmer is not noticeably more effective than a programmer armed with an ordinary optimizing compiler
- Second epoch: The AI can understand a small handful of components and make improvements to them, but the total amount of AI-driven improvement is small by comparison with programmer-driven development. Sufficiently major programmer improvements do very occasionally trigger secondary improvements. The total amount of work done by the AI on its own subsystems serves only as a measurement of progress and does not significantly accelerate work on AI programming
- Third epoch: AI-driven improvement is significant, but development is "strongly" programmer-dominated in the sense that overall systemic progress is driven almost entirely by the creativity of the programmers. The AI may have taken over some significant portion of the work from the programmers. The AI's domain competencies for programming may play a critical role in the AI's continued functioning
- Fourth epoch: AI-driven improvement is significant, but development is "weakly" programmer-dominated. AI-driven improvements and programmer-driven improvements are of roughly the same kind, but the programmers are beter at it. Alternatively, the programmers have more subjective time in which to make improvements, due to the number of programmers or the slowness of the AI
- Epochs for overall intelligence:
    - Tool-level AI: The AI's behaviors are immediately and directly specified by the programmers, or the AI "learns" in a single domain using prespecified learning algorithms
    - Prehuman AI: The AI's intelligence is not a significant subset of human intelligence. Nonetheless, the AI is a cognitive supersystem, with some subsystems we would recognize, and at least some mind-like behaviors.
    - Infrahuman AI: The AI's intelligence is, overall, of the same basic character as human intelligence, but substantially inferior. The Ai may excel in a few domains where it possesses new sensory modalities or other brainware advantages not available to humans.

3 See also

4 References

- [Levels of Organization in General Intelligence](#)