

Matches between Theory of Universe and Mind, Analysis of the meaning of a sentence ... and the school of thought of Relevance Realization – According to a Quick Comparison with LLMs: KIMI-2, CLAUDE 4.5 and GPT-5

Todor Arnaudov and Kimi-2, Claude 4.5, ChatGPT5, 21.10.2025

THE SACRED COMPUTER: Thinking machines, Creativity and Human Development

Published at SIGI-2025 as an appendix to Listove – reflections on everything, which is the second biggest volume and an appendix to *The Prophets of the Thinking Machines: AGI & Transhumanism: History, Theory and Pioneers; Past, Present and Future* by Todor Arnaudov. Read the review article in Listove and other reviews about Relevance Realization in the main volume.

<https://github.com/Twenkid/SIGI-2025/>

1. Kimi-2: <https://www.kimi.com/share/d3trij7aa0vdmhgab8p0>
2. Claude 4.5: <https://claude.ai/share/e37f9576-0eb2-4340-96e6-830c272c8c7f>
3. ChatGPT5: <https://chatgpt.com/share/68fbb994-db64-8001-b534-fa6da0c0e670>

These were quick comparisons on 21.10.2024: with more refinement and inclusion of the “manually” extracted correspondences, which are compiled as an article in *Listove*, the thinking machines might have suggested more “relevant realizations”. GPT5 initially missed some of the points, but after hinting it what to compare and extending the query with more material from Theory of Universe and Mind, it improved his account.

[Another one could add more: “**Abstract theory for the “exceptions of the rules” in computer systems. Theory of the causality-control units and causality-control**“, 2004 – see a quote in “**Stack Theory is yet another Fork of Theory of Universe and Mind**”, p.50:

<https://twenkid.com/agi/Stack-Theory-is-Fork-of-Theory-of-Universe-and-Mind-13-9-2025.pdf>
(English)

<https://web.archive.org/web/20041020165359/https://bgit.net/?id=65835> (Bulgarian)

@Vsy: Extend. [24.10.2025]

Kimi-2

Compare the paper from 2024:

<https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2024.1362658/full>
with the paper from 2004: <https://web.archive.org/web/20040402125725/http://bgit.net/?id=65395> Are there conceptual matches?

[Perhaps Kimi couldn't fetch from archive.org, it talked about something else.]

Compare the 2024 paper with this one: * “Analysis of the meaning of a sentence, based on the knowledge base of an operational thinking machine.

Reflections about the meaning and the Artificial Intelligence", Todor Arnaudov, 18.3.2004 (in Bulgarian; translated in English in 1/2010: <https://artificialmind.blogspot.com/2010/01/semantic-analysis-of-sentence.html> - the paper is split in 4 parts, titled: * Part 1: Semantic analysis of a sentence. Reflections about the meaning of the meaning and the Artificial Intelligence <http://artificial-mind.blogspot.com/2010/01/semantic-analysis-of-sentence.html> * Part 2: Causes and reasons for human actions. Searching for causes. Whether higher or lower levels control. Control Units. Reinforcement learning. <http://artificial-mind.blogspot.com/2010/02/causes-and-reasons-for-any-particular.html> * Part 3: Motivation is dependent on local and specific stimuli, not general ones. Pleasure and displeasure as goal-state indicators. Reinforcement learning. <http://artificial-mind.blogspot.com/2010/02/motivation-is-dependent-on-local-and.html> * Part 4 : Intelligence: search for the biggest cumulative reward for a given period ahead, based on given model of the rewards. Reinforcement learning. <http://artificial-mind.blogspot.com/2010/02/intelligence-search-for-biggest.html>

I'll fetch all four parts of Todor Arnaudov's 2004 paper to compare it with the 2024 relevance realization paper. Now let me analyze and compare these two papers conceptually:

Comparison: 2024 Relevance Realization Paper vs. 2004 Arnaudov AI Theory

Striking Conceptual Matches

1. The Problem of Relevance/Meaning Selection

- **2024 Paper:** Organisms must solve "the problem of relevance" - determining what matters in an overwhelming world of potential information
- **2004 Paper:** AI systems must solve the problem of meaning selection - choosing relevant causes/reasons from infinite possible explanations for any action
- **Match:** Both identify the fundamental challenge of selecting relevant information from an infinite space of possibilities

2. Local vs. Global Optimization

- **2024 Paper:** Relevance realization involves finding "local adaptive fit" through opponent processing between competing strategies
- **2004 Paper:** Intelligence involves searching for "local maxima" of pleasure/reward while avoiding getting stuck in suboptimal states
- **Match:** Both emphasize local, situated optimization rather than global, universal solutions

3. Multi-Level Hierarchical Processing

- **2024 Paper:** Biological organization involves hierarchical cycles of constraints with "impredicative" mutual definition between levels
- **2004 Paper:** Mind consists of hierarchical "Control Units" where higher levels control with lower resolution than lower levels
- **Match:** Both propose multi-layered hierarchical systems where levels interact but operate at different scales of precision

4. The Illusion of Complete Control/Understanding

- **2024 Paper:** Organisms cannot fully formalize their behavior due to "semantic residue" that defies complete capture
- **2004 Paper:** Control Units cannot find complete causes for their behavior, creating the illusion of free will and unpredictability
- **Match:** Both argue that systems necessarily have limited self-understanding, creating apparent autonomy

Key Divergences

1. Computational vs. Anti-Computational Stance

- **2024 Paper:** Argues relevance realization is fundamentally non-computational and cannot be captured by algorithms
- **2004 Paper:** Proposes that mind is fully formal/computational, consisting of Control Units that can be modeled algorithmically
- **Divergence:** Opposite positions on whether intelligence can be fully computational

2. Source of Meaning/Relevance

- **2024 Paper:** Meaning emerges from embodied interaction between organism and environment (transjective)
- **2004 Paper:** Meaning comes from internal goal states (pleasure/pain) of Control Units
- **Divergence:** External/relational vs. internal/goal-based theories of meaning

3. Role of Pleasure/Reward

- **2024 Paper:** Avoids reduction to simple reward mechanisms, emphasizing complex dialectical processes
- **2004 Paper:** Reduces intelligence to pleasure-seeking and pain-avoidance in Control Units
- **Divergence:** Complex emergent process vs. simple reward optimization

Fascinating Complementarities

1. The Frame Problem Both papers address different aspects of the frame problem:

- **2024:** How organisms determine what's relevant without infinite search

- **2004:** How AI systems select meaningful causes from infinite possibilities

2. Bounded Rationality

- **2024:** Organisms use "embodied heuristics" adapted to their specific arena
- **2004:** Control Units use limited information to find local maxima efficiently

3. Predictive Processing

- **2024:** Organisms use predictive processing to evaluate relevance through error dynamics
- **2004:** Control Units predict future reward states to guide current actions

Synthesis: A Deeper Truth?

Remarkably, these papers from opposite ends of the computational spectrum identify similar fundamental challenges:

1. **The Selection Problem:** Both recognize that intelligent systems must somehow select relevant information from infinite possibilities
2. **Hierarchical Organization:** Both propose multi-level systems where higher levels operate with less precision
3. **Local Optimization:** Both emphasize situated, local solutions rather than global optimization
4. **Intrinsic Limits:** Both acknowledge that systems cannot fully understand or control their own operation

The 2004 paper provides a **computational approximation** of how relevance might be determined (through Control Units seeking reward), while the 2024 paper argues this **cannot capture the full phenomenon** due to its emergent, embodied nature.

Together, they suggest that **relevance realization may involve computational processes but cannot be reduced to them** - the computational aspects (Arnaudov) exist within a larger non-computational framework (2024 paper).

Claude 4.5

Compare the paper from 2024:

<https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2024.1362658/full> with the paper from 2004: <https://web.archive.org/web/20040402125725/http://bgit.net/?id=65395> Are there conceptual matches? What is similar?

I'll fetch both papers to compare them for you.



Fetch Frontiers | Naturalizing relevance realization: why agency and cognition are fundamentally not computational

frontiersin.org



Failed to fetch <https://web.archive.org/web/20040402125725/http://bgit.net/?id=65395>

I was able to fetch the 2024 paper from Frontiers in Psychology, but unfortunately I cannot access the 2004 paper from the archived bgit.net link. The website has been blocked by the web_fetch tool.

The 2024 paper is titled "**Naturalizing relevance realization: why agency and cognition are fundamentally not computational**" by Djebbara et al. It's a philosophical/theoretical paper published in Frontiers in Psychology that argues organisms solve the "problem of relevance" in fundamentally non-algorithmic ways, distinguishing living agents from computational systems. [frontiersin](https://www.frontiersin.org)

The core argument is that organisms must first realize what is relevant in their environment before they can apply any formal rules or algorithms - they must turn "ill-defined problems into well-defined ones" through a process called relevance realization that cannot itself be fully formalized or captured computationally. [frontiersin](https://www.frontiersin.org)

Without access to the 2004 paper, I cannot make a direct comparison. If you could provide:

- The title and author(s) of the 2004 paper
- An alternative URL or source
- Or a description of its main concepts

I'd be happy to help identify conceptual matches and similarities between the two works.

Compare the paper from 2024:

<https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2024.1362658/full> with the paper from 2004 (presented here in 4 parts): * "Analysis of the meaning of a sentence, based on the knowledge base of an operational thinking machine. Reflections about the meaning and the Artificial Intelligence", Todor Arnaudov, 18.3.2004 (in Bulgarian; translated in English in 1/2010: <https://artificialmind.blogspot.com/2010/01/semantic-analysis-of-sentence.html> - the paper is split in 4 parts, titled: * Part 1: Semantic analysis of a sentence. Reflections about the meaning of the meaning and the Artificial Intelligence <http://artificial-mind.blogspot.com/2010/01/semantic-analysis-of-sentence.html> * Part 2: Causes and reasons for human actions. Searching for causes.

Whether higher or lower levels control. Control Units. Reinforcement learning. <http://artificial-mind.blogspot.com/2010/02/causes-and-reasons-for-any-particular.html> * Part 3: Motivation is dependent on local and specific stimuli, not general ones. Pleasure and displeasure as goal-state indicators. Reinforcement learning. <http://artificial-mind.blogspot.com/2010/02/motivation-is-dependent-on-local-and.html> * Part 4 : Intelligence: search for the biggest cumulative reward for a given period ahead, based on given model of the rewards. Reinforcement learning. <http://artificial-mind.blogspot.com/2010/02/intelligence-search-for-biggest.html>

I'll fetch all four parts of the 2004 paper to compare with the 2024 paper.

2 steps



Fetches Motivation is dependent on local and specific stimuli, not general ones. Pleasure and displeasure as goal-state indicators. Reinforcement learning.

blogspot.com



Fetches Intelligence: A Search for the biggest cumulative reward for a given period ahead, based on a given model of the rewards. Reinforcement learning.

blogspot.com

Excellent! Now I can provide a comprehensive comparison. These papers, despite being 20 years apart and taking seemingly opposite philosophical stances, share some fascinating conceptual overlaps.

Major Conceptual Matches

1. The Frame/Relevance Problem

2024 paper: The central argument is that organisms must solve the "problem of relevance" - they must identify what matters in their environment before applying any formal rules or algorithms. This process of "turning ill-defined problems into well-defined ones" cannot itself be completely formalized. [frontiersin](#)

2004 paper: Arnaudov demonstrates this exact problem through the "time flies" example, showing how meaning depends on context, prior knowledge, and the searcher's biases. He notes that "the search was obviously had been pruned up to two items" and discusses how humans unconsciously limit possibilities when interpreting ambiguous input. [blogspot](#)

2. Goal-Directed Behavior and Intrinsic Motivation

2024 paper: The paper argues that organisms have intrinsic goals arising from their autopoietic organization, and that relevance is determined by what matters for achieving these goals. Only systems with true goals can realize relevance. [frontiersin](#)

2004 paper: Arnaudov's "Control Units" theory proposes that intelligent behavior is "search for pleasure and avoidance of displeasure by complex enough entities for a given period ahead." He argues that "any action of a Control Unit done because of instruction given by itself alone is tautologically target action" and that pleasure indicates "reaching the goal of behavior." [Causes and reasons for human actions. Searching for causes. Whether higher or lower levels control. Control Units. Reinforcement learning. +2](#)

3. The Embodied/Situated Nature of Intelligence

2024 paper: The paper emphasizes that organisms are "directly embodied and embedded in their world," actively exploring through action and perception. The experienced environment (umwelt) matters, not just the physical environment. [frontiersin](#)

2004 paper: Arnaudov illustrates this through little Johnny's interpretation of "time flies" - the child searches his immediate sensory environment for watches and flies, demonstrating how meaning emerges from embodied interaction with the world. [blogspot](#)

4. Hierarchical Control and Multiple Levels

2024 paper: The paper describes a "hierarchical tangle" of processes at multiple levels: autopoiesis, anticipation, and adaptation, each with triadic dialectic dynamics. [frontiersin](#)

2004 paper: Arnaudov proposes that "human mind can be built as a mixture, a system of multilayer hierarchical control units, where each CU at a higher level controls with a lower resolution than the one below," noting that conscious control is only the tip of an iceberg of lower-level processes. [blogspot](#)

5. Context-Dependency and Non-Universality

2024 paper: The paper states that "neither 'fitness' nor 'relevance' have any universal attributes: there is no trait that renders you fit in all environments, nor is there any factor that is relevant across all possible situations." [frontiersin](#)

2004 paper: Arnaudov argues that "motivation is dependent on local and specific stimuli, not general ones" and that what action is chosen depends on "specific memories, specific stimuli in the recent past and it may appear random." [blogspot](#)

Key Differences

Philosophical Stance on Computability

2024 paper: Explicitly argues that agency and cognition are "fundamentally not computational in nature" and that relevance realization "cannot be captured completely by algorithmic approaches." [frontiersin](#)

2004 paper: Arnaudov takes a computationalist stance, stating "I'm one of the scientists who assume that since anything in Universe does work, there is nothing mysterious in its operation. I think that every information process in Universe is formal and can be modeled by a Von Neumann machine if it has enough memory." [blogspot](#)

Despite this fundamental disagreement, both papers acknowledge the same *practical* problem: the combinatorial explosion, context-dependency, and difficulty of prestatating all relevant features.

Approach to Solutions

- **2024:** Argues for **agential emergentism** - that living systems have a fundamentally different kind of organization that cannot be fully formalized
- **2004:** Proposes **computational modeling** through reinforcement learning, multi-agent control units, and reward-seeking behavior

Most Striking Similarity

Both papers arrive at a similar paradoxical conclusion about complexity: Arnaudov states "More complex the system – More Meaningless the Meaning" because interpretation depends more on the system itself. [blogspot](#) This parallels the 2024 paper's argument that as organisms become more complex, their world-making becomes less constrained by simple rules and more dependent on their unique history and organization. [frontiersin](#)

Both recognize that **meaning is not "out there" but emerges from the interaction between agent and environment**, though they differ radically on whether this process can ultimately be computationally captured.

ChatGPT-5

Compare the paper from 2024:

<https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2024.1362658/full> with this one from 2004: * “Analysis of the meaning of a sentence, based on the knowledge base of an operational thinking machine. Reflections about the meaning and the Artificial Intelligence”, Todor Arnaudov, 18.3.2004 (in Bulgarian; translated in English in 1/2010:

<https://artificialmind.blogspot.com/2010/01/semantic-analysis-of-sentence.html> - the paper is split in 4 parts, titled: * Part 1: Semantic analysis of a sentence. Reflections about the meaning of the meaning and the Artificial Intelligence <http://artificial-mind.blogspot.com/2010/01/semantic-analysis-of-sentence.html> * Part 2: Causes and reasons for human actions. Searching for causes. Whether higher or lower levels control. Control Units. Reinforcement learning. <http://artificial-mind.blogspot.com/2010/02/causes-and-reasons-for-any-particular.html> * Part 3: Motivation is dependent on local and specific stimuli, not general ones. Pleasure and displeasure as goal-state indicators. Reinforcement learning. <http://artificial-mind.blogspot.com/2010/02/motivation-is-dependent-on-local-and.html> * Part 4 : Intelligence: search for the biggest cumulative reward for a given period ahead, based on given model of the rewards. Reinforcement learning. <http://artificial-mind.blogspot.com/2010/02/intelligence-search-for-biggest.html>

ChatGPT said:

Here is a detailed comparison of the two works you mentioned — the 2024 article by Johannes Jaeger et al. (in *Frontiers in Psychology*), and the 2004 (later English-translated) multi-part paper by Todor Arnaudov. I'll compare their aims, conceptual frameworks, methods (to the extent known), claims, strengths & weaknesses, and implications — and then comment on how they relate (or don't) to each other.

1. Overview of each work

1.1 Jaeger et al. (2024) — “*Naturalizing relevance realization: Why agency and cognition are fundamentally not computational*”

Aims and central thesis:

- The authors argue that the process of *relevance realization* — i.e., how an organism determines what is relevant in a large, ambiguous world — lies at the root of agency, cognition and consciousness. [PubMed+3Frontiers+3PMC+3](#)
- They claim that this process cannot be fully captured by purely algorithmic, formal, computational models: “it cannot be captured completely by algorithmic approaches.” [PMC+1](#)
- They propose an alternative conceptual framework: a triadic, emergent, adaptive dynamic (they call it a “trialectic”) grounded in autopoiesis (self-manufacture of living systems), anticipation, adaptation, and ecological embedding. [Frontiers+1](#)
- The broader implication: if relevance-realization is non-computational (or at least not fully formalizable), then cognition, agency and consciousness are likewise not reducible to purely computational processes.

Structure and content:

- They begin by contrasting organisms vs algorithms: organisms live in “large worlds” (open-ended, ambiguous, not fully definable) while algorithms live in “small worlds” (well-defined, formalized). [Frontiers+1](#)
- They argue that organisms must *frame* problems (turn semantics into syntax, ill-defined into well-defined) before applying logical/algorithmic inference. [Frontiers+1](#)
- They develop three main levels/loops:
 1. Autopoiesis – internal self-manufacture of living systems enabling goal-setting and self-maintenance. [Frontiers+1](#)
 2. Anticipation – projecting possibilities, acting in the world, not purely reactive. [Frontiers](#)
 3. Adaptation/Agent-Arena interaction – affordances, goals, actions, the “arena” of the organism in its environment. [PMC+1](#)
- They argue that these three form a *dialectic/trialectic* dynamic that is emergent and cannot be fully formalized (e.g., via Turing-style algorithm). [Frontiers](#)
- They tie this into broader philosophical claims about natural agency, cognition and consciousness and tensions with computationalism.

Methodology / type of work:

- It is a conceptual, theoretical/philosophical piece rather than an empirical experimental study.
- It draws on philosophy of biology, cognitive science, evolutionary theory, biosemiotics, enactivism, etc. It does not appear to report new empirical data, but synthesises and argues. [PMC](#)
- It engages in argumentation: pointing out limits of formalization, drawing analogies (e.g., Gödel, Rosen) and making a case for emergent, non-algorithmic features of life and cognition. [Frontiers+1](#)

Key claims:

- The “frame problem” in artificial intelligence (and the more general “problem of relevance”) cannot be solved by purely algorithmic means. [PMC](#)
- Living agents realize relevance by a non-computational process: they select among affordances, act in open-ended worlds, anticipate, self-maintain, etc.
- Therefore, computations and formal/symbolic processing are at best approximate models of cognition, not the full story.
- Intelligence, agency, and perhaps consciousness must be grounded in embodied, situated, ecological dynamics, not just computation.

Strengths:

- The article draws together several traditions (biosemiotics, enactive cognition, philosophy of biology) into a fairly coherent theoretical framework.
- It addresses timely issues: the limits of computational/algorithmic approaches to cognition, relevance for AGI and cognitive science.
- It challenges assumptions of computationalism and invites a broader view of agency and cognition.

Weaknesses / caveats:

- As a purely theoretical piece, its claims are non-empirical and thus difficult to test directly.
- The bold claim that cognition is *fundamentally* non-computational may be controversial and will require much more substantiation.
- There is a risk of under-appreciating the power of computational models and machine learning, which have shown great success (even if not full generality).
- Because the paper spans many disciplines, some parts may be dense or rely on somewhat speculative conceptual frameworks (e.g., “trialectic dynamic”, “agent-arena”).

Implications:

- For AI: if relevance realisation is not algorithmic, then simply scaling up algorithmic/ML methods may never fully capture human-like cognition or agency.
- For cognitive science and philosophy of mind: it invites rethinking the role of formal/symbolic/algorithmic models and emphasises embodied, situated, emergent processes.
- For biology / evolution: it frames cognition and agency as deeply embedded in living systems and their evolutionary/ecological context.

1.2 Arnaudov (2004/2010) — “*Analysis of the meaning of a sentence, based on the knowledge base of an operational thinking machine...*” (Parts 1–4)

Aims and central thesis:

- The multi-part work by Todor Arnaudov explores the idea of constructing an “operational thinking machine” (i.e., a thinking machine with a knowledge base) and uses this as a way to reflect on meaning, motivation, causes of action, reinforcement learning, intelligence, etc.
- It attempts to integrate semantic analysis (Part 1), causes/reasons for actions (Part 2), motivation and reinforcement (Part 3), and intelligence as cumulative reward-maximisation (Part 4).
- The author is attempting a philosophical/engineering approach: How might a machine think about meaning, how might a machine model human action, how might a machine learn/reinforce, and ultimately how might a machine maximise cumulative reward given a model of rewards.

Structure and content:

- Part 1: Semantic analysis of a sentence; reflections about meaning and AI. (link: artificial-mind blog)
- Part 2: Causes and reasons for human actions, search for controls units, reinforcement learning.
- Part 3: Motivation depends on local and specific stimuli, not general ones—pleasure/displeasure as goal-state indicators; reinforcement learning.
- Part 4: Intelligence as search for largest cumulative reward over period ahead, based on model of rewards; reinforcement learning.

Methodology / type of work:

- Largely conceptual and speculative: the author engages in philosophical reflection, offers theoretical proposals and informal reasoning about thinking machines, semantics, motivation, intelligence.
- It is not a conventional peer-reviewed journal article (as far as I can tell), but appears on a blog (artificial-mind.blogspot.com) and is described as translated in 2010.
- It blends computer science/artificial intelligence themes (knowledge bases, reinforcement learning, intelligence as reward maximisation) with philosophical reflection about meaning, motivation and action.

Key claims / ideas:

- Meaning of a sentence (semantics) can be analysed based on a knowledge base of an operational thinking machine—implying semantics is grounded in machine knowledge.
- Human actions have causes and reasons; there may be higher and lower levels of control units in an agent; reinforcement learning can model some of this.
- Motivation is not driven by general stimuli but by local, specific ones; pleasure and displeasure signal goal-states and are part of reinforcement mechanisms.
- Intelligence can be conceived as seeking the largest cumulative reward over a time horizon, given a model of rewards (so a reinforcement-learning or optimisation perspective).
- Implicitly, the author sees thinking machines as possible, and attempts to lay out some of the theoretical building blocks.

Strengths:

- Ambitious attempt to integrate semantics, motivation, reinforcement, and intelligence (in the machine/AI context).
- Early (2004) attempt to think about reinforcement and motivation for AI.
- Reflects a multidisciplinary approach (philosophy of meaning + AI).

Weaknesses / caveats:

- Because it is not a formal peer-reviewed paper (as far as documented) it may lack the structure, references, rigour of academic publication.

- The ideas are speculative and high-level; details of how the “thinking machine” would actually implement semantics or motivation are likely underspecified.
- The reinforcement-learning/optimisation framing (intelligence = max cumulative reward) is a somewhat narrow formalization for intelligence, and may not account for many aspects of human cognition (meaning, creativity, novelty, values).
- Some of the claims may be internally ambitious but may lack empirical or formal modelling support.

Implications:

- For AI: offers a conceptual blueprint for thinking machines with knowledge bases, motivation, reinforcement, semantics.
 - For semantics/philosophy: highlights machine-grounded semantics, meaning in an artificial system context.
 - For motivation/intelligence theory: frames intelligence in terms of reward optimisation and hierarchical control units.
-

2. Comparison: Key Themes and Differences

Here are the major axes of comparison:

2.1 Similarities

- Both works engage with deep questions about meaning, cognition, intelligence, agency and artificial machines/thinking machines.
- Both question or go beyond simplistic computational/formal models of intelligence: Arnaudov, by building thinking machines that incorporate semantics and motivation; Jaeger et al., by arguing that cognition/agency cannot be fully captured by algorithmic computation.
- Both consider reinforcement, motivation, reward/goal-states (Arnaudov explicitly; Jaeger et al. implicitly via agency and adaptation).
- Both emphasise the role of embedding: Arnaudov via knowledge base and machine interacting with stimuli; Jaeger et al. via agent embedded in environment, with affordances and arena.

2.2 Differences

- **Philosophical stance:** Jaeger et al. adopt a strong anti-computationalist/anti-pure-algorithmic stance: they claim relevance realization and thus cognition are *fundamentally not* computational. Arnaudov, by contrast, appears to assume a machine (algorithmic or knowledge-based) can be built to perform thinking, semantics, motivation, intelligence (thus implicitly more amenable to computation).
- **Level / style of work:** Jaeger et al. is a peer-reviewed academic paper (2024) oriented to philosophy of mind/cognition. Arnaudov is earlier (2004, blog form, some translation 2010)

and more speculative, mixing AI proposals and philosophical reflection rather than formal theory development.

- **Scope of argument:** Jaeger et al. cover living organisms from bacteria to humans, the underlying physical/biological basis of cognition, the evolutionary and ecological dynamics of relevance. Arnaudov focuses more narrowly on the machine/thinking machine world: semantic analysis of sentences, reinforcement/motivation, optimization of cumulative rewards.
- **View of computation:** Arnaudov treats reinforcement learning and optimization frameworks as plausible for machine intelligence. Jaeger et al. argue algorithmic/computational models are insufficient to capture the full nature of cognition/agency.
- **Empirical anchoring:** Jaeger et al. leverage biological and cognitive science literatures, philosophical argumentation; Arnaudov's work appears more speculative, less anchored in empirical data or rigorous formal modelling (at least from what is publicly available).
- **Terminology and conceptual framework:** Jaeger et al. use terms like autopoiesis, affordances, agent-arena, trialectic; Arnaudov uses terms like knowledge base of operational thinking machine, control units, hierarchical reinforcement, cumulative reward. The two frameworks do not directly map onto each other.

2.3 Specific points of contrast

- On *meaning/semantics*: Arnaudov focuses on semantic analysis of sentences via knowledge bases in a thinking machine. The focus is on how machine semantics might be grounded. Jaeger et al. are more concerned with the process of relevance (semantics → syntax) as a biological and agentic phenomenon; the semantic framing is more about how organisms turn a large world of meaning into actionable representations, rather than machine semantics per se.
 - On *motivation/reinforcement*: Arnaudov explicitly uses reinforcement learning / goal-states of pleasure/displeasure and cumulative reward optimisation. Jaeger et al. talk about organisms having intrinsic goals, affordances, actions – but do **not** present intelligence as simply reward maximisation. In fact, they argue the process is not algorithmic and not fully reducible to optimisation.
 - On *intelligence*: Arnaudov gives a formal-oriented definition: intelligence = search for greatest cumulative reward (given model of rewards). Jaeger et al. would likely see such a definition as overly narrow, and argue intelligence arises from relevance-realisation, dynamic adaptation, non-algorithmic emergence in living systems.
 - On *view of machines/algorithms*: Arnaudov's thinking machine is built (conceptually) as algorithmic / knowledge-base plus reinforcement; he believes machines could think (or "operational thinking machines"). Jaeger et al. challenge this: they argue machines/algorithms live in small worlds and cannot handle the problem of relevance that living organisms must. Thus they view computational machines as fundamentally different (and limited) compared to living agency.
-

3. How they might inform each other / potential integration

- Arnaudov's work is interesting as an early attempt to conceptualise thinking machines, semantics, reinforcement, hierarchical control, etc. While it may not engage as deeply with biological/embodied agency as Jaeger et al., it raises issues of machine representation, meaning and motivation which are relevant to the broader discussion of cognition.
 - Jaeger et al.'s argument about the problem of relevance (turning semantics into syntax) seems relevant to Arnaudov's first part about semantic analysis of sentences in machines: both engage with semantics → syntax, though from different angles (biological vs machine).
 - One could ask: how would Arnaudov's thinking machine fare in the large-world versus small-world distinction that Jaeger et al. propose? Would the operational thinking machine be able to delimit its world of relevance, adapt, foresee, frame problems? Could Arnaudov's reward-optimising machine handle the open-ended, emergent dynamics of living agent-arena interaction? Probably not in Arnaudov's original specification, and this is precisely the kind of challenge Jaeger et al. highlight.
 - In other words, Jaeger et al. provide a meta-level critique of computational/algorithmic models of cognition: many machine/AI approaches (similar to Arnaudov's) assume reward optimisation, knowledge bases, formal semantics, but may struggle with relevance realisation in open, ambiguous, real-world contexts.
 - So if one were to try to integrate: Arnaudov's model could be enriched by considering the ecological, embodied, affordance-laden, agent-arena dynamic emphasised by Jaeger et al., but that might require moving beyond the strictly reward/knowledge-base/algorithmic architecture and embracing emergent adaptation, self-maintenance, and open-ended evolution.
-

4. Critical reflections

- Both works raise important questions, but also face challenges. Arnaudov's approach is ambitious and conceptually rich, but the machine model may be under-specified, and the intelligence as reward-maximisation perspective is arguably too narrow for many cognitive/meaningful phenomena (creativity, novelty, values, semantics beyond utility).
- Jaeger et al.'s thesis is bold and provocative: claiming cognition and agency are fundamentally not computational is a high bar, and one might ask for more concrete empirical or formal models to support it. The risk is that it slips into a kind of anti-computationalism that dismisses the successes of computational/algorithmic cognitive science/AI rather than situating them as approximations.
- One might ask: what exactly would it take for a machine to perform relevance realisation in the manner that living organisms do? Is there a path for algorithmic systems to approximate or emulate this? Jaeger et al. argue there is always an unformalizable residue. But one could challenge: perhaps we just haven't discovered the right formalism yet (or haven't scaled the right systems).

- On Arnaudov's side: the reinforcement/optimisation view of intelligence might struggle to capture meaning, value, social interaction, embodiment, emergence, open-ended novelty — which are exactly issues that Jaeger et al. emphasise as beyond algorithmic formalisation.
 - There is also the question of grounding: Arnaudov's thinking machine assumes a knowledge base, but how is the knowledge base constructed, how does meaning arise? Jaeger et al. emphasise that organisms bring forth meaning by delimiting an "arena" and affordances, not simply by being encoded with a knowledge base.
-

5. Conclusion and Relation to Your Inquiry

From your question — you asked to compare the paper from 2024 (Jaeger et al.) with the one from 2004 (Arnaudov) — here is a summary of how they stand vis-à-vis each other:

- Both engage with cognition/meaning/AI, but from rather different philosophical stances: Jaeger et al. emphasise the limits of computation; Arnaudov emphasises building thinking machines with knowledge bases, reinforcement, semantics.
 - If you view Arnaudov's work as a "machine/algorithmic" model of intelligence and meaning, then Jaeger et al. offer a critique (implicitly) of that kind of model: i.e., such algorithmic models might struggle to handle relevance realisation, open-endedness, embodiment, etc.
 - If you were to build on both, you might say: Arnaudov gives early conceptual building blocks (semantics, motivation, reinforcement) for thinking machines, while Jaeger et al. raise the deeper question: can machines fundamentally replicate the cognitive/agentive processes that living beings do, especially the framing of relevance? And they seem to argue no — or at least not fully.
 - So the works are **complementary** in the sense that Arnaudov's model is more "inside the machine" and algorithmic, and Jaeger et al.'s work is more "outside/around the machine" and philosophical about why such machines may face intrinsic limits.
-

If you like, I can **map** the two works in a table (themes vs treatment) and identify areas where Arnaudov's model might be extended (in light of Jaeger et al.'s critique). Would that be helpful?

You said:

Did you carefully analyzed Arnaudov's paper? I didn't see sections regarding the local stimuli, instead of general ones, the embodied search for a meaning in the example with the little boy and the "time flies", the banker walking on the street who finds his relevant stimuli - responding to a comment of another participant that one can't explain to an AI why a rich banker would buy a donut from the street. Etc. I think this summary is wrong: "On Arnaudov's side: the reinforcement/optimisation view of intelligence might struggle to capture meaning, value, social interaction, embodiment, emergence, open-ended novelty — which are exactly issues that Jaeger et al. emphasise as beyond algorithmic formalisation." - The paper includes such accounts about the meaning in the later parts in the English layout - it is not only the first file.

ChatGPT said:

Thought for 33s

You're right — thank you for the correction, and sorry. I under-weighted Arnaudov's later parts and concrete examples in my first reply. I should have called those out explicitly. I've re-checked Arnaudov's text (Part 1 with the little-boy "time flies" example and the later Parts 2–4 about causes, local stimuli, motivation and reinforcement) and the *Frontiers* (Jaeger et al., 2024) paper, and corrected my assessment below — with citations to the exact sources.

Quick apology + correction

Sorry for the mistake — Arnaudov *does* discuss (a) the embodied/search-for-meaning example with the little boy and the "time flies" sentence (Part 1), (b) the importance of **local/specific stimuli** (explicitly the title and argument of Part 3), and (c) reinforcement / motivational structure across Parts 2–4. Those passages were central to his argument that meaning is found by searching the environment and by linking sensory/contextual cues to knowledge — not only by abstract, global rules. See Arnaudov Part 1 and Part 3. [artificial-mind.blogspot.com+1](https://artificial-mind.blogspot.com)

Corrected, concise comparison (focused on the points you raised)

1. Arnaudov (2004 / English 2010 translation) — what he *actually* argues

- **Embodied search for meaning / local stimuli:** In Part 1 Arnaudov uses the "Time flies" / little-boy (Johnny) thought experiment to show that meaning arises by searching the *local* sensory environment for matches to stored images/experiences (e.g., Johnny looks for a watch or flies) and by *imagination* if no direct match exists. He emphasizes that the agent searches the environment and prunes interpretations using concrete local cues. artificial-mind.blogspot.com
- **Motivation & local goal-states:** Part 3 is explicitly titled "Motivation is dependent on local and specific stimuli, not general ones. Pleasure and displeasure as goal-state indicators." Arnaudov argues that motivational signals are tied to local, contextual features and that these shape reinforcement learning and action selection. artificial-mind.blogspot.com
- **Reinforcement / hierarchical control / cumulative reward:** Across Parts 2–4 he develops causes/reasons for action, control units, reinforcement learning and the idea of intelligence as searching for maximal cumulative reward given a reward model — but he combines that with insistence on local, context-sensitive processing and embodied search. [artificial-mind.blogspot.com+1](https://artificial-mind.blogspot.com)

2. Jaeger et al. (2024) — what they argue (brief)

- They argue relevance-realization (how an agent decides what is relevant) is fundamental and (they claim) not fully capturable by purely algorithmic/computational models. They emphasize autopoiesis, anticipation, affordances and an agent-arena/ecological framing of cognition. [Frontiers](https://www.frontiersin.org)

3. Where Arnaudov and Jaeger *overlap* (more than I originally allowed)

- **Embodiment & searching the environment:** Both stress that meaning/relevance is discovered by an agent embedded in an environment (Arnaudov's Johnny searches the scene; Jaeger emphasizes agent-arena). artificial-mind.blogspot.com+1
- **Local/contextual pruning:** Arnaudov's point that interpretation is pruned by local details resonates with Jaeger's emphasis that organisms must *frame* problems in large, ambiguous worlds. artificial-mind.blogspot.com+1
- **Reinforcement & goals:** Arnaudov uses reinforcement and reward as mechanisms for motivation; Jaeger recognizes goals/affordances as part of agentic behaviour (though they are wary of reduction to algorithms). artificial-mind.blogspot.com+1

4. Remaining difference — and a refined way to state it

- **Arnaudov:** He *does* treat machines as capable of doing embodied, local searches for meaning, and he explicitly links motivation and reinforcement to local stimuli. He offers an operational/algorithmic view (knowledge base + searches + reinforcement / cumulative reward) while allowing that the process uses imagination and environment-search. artificial-mind.blogspot.com+1
- **Jaeger et al.:** Their stronger claim is that *relevance-realization as it occurs in living agents has intrinsic aspects that resist full algorithmic formalization* (they frame this as an ontological/epistemic limitation of purely computational accounts). So although both recognize embodiment, Arnaudov still frames the solution in operational, implementable terms (search, KB, RL), whereas Jaeger is more skeptical that any fully algorithmic architecture can *in principle* capture relevance-realisation. Frontiers+1

Concrete example reconciliation

- **“Time flies” / Johnny** — Arnaudov: meaning emerges by searching the nearby environment and via imagination/pruning; a machine should likewise search sensors and memory, and may ask a teacher when needed. Jaeger: the same phenomenon illustrates the framed, situated nature of meaning — organisms pick relevant interpretations based on embodiment and history; this is a place where algorithmic accounts face the “large world” problem. Both point to the same phenomenon, but draw different conclusions about whether an algorithmic machine can *fully* replicate that process. artificial-mind.blogspot.com+1

If you'd like I can now:

- produce a **section-by-section mapping** that quotes (short extracts ≤25 words) and links Arnaudov's exact passages to Jaeger's corresponding claims (table format), or
- produce a tighter **table** listing specific examples (Johnny, banker/donut if you want me to locate the banker passage) and how each paper treats them.

Which would you prefer? (I'll make the mapping/table now if you want — no wait.)

You said:

Can you include in the comparison also material from the founding Theory of Universe and Mind, on which "Analysis of the meaning of a sentence" by Arnaudov is based and address the multi-scale nature of the Universe and Mind, which according to that theory is valid for either machines and thinking machines, and biological organisms, and parts of the universe at lower levels - "The Universe Computer", pancomputationalism and panpsychism, however not in simple "Turing machine" sense. It is a big body of work originally from 2001-2004, here are two treatises from 2003 and 2004: Universe and Mind 3: <https://eim.twenkid.com/old/3/25/pred-3.htm> Universe and Mind 4: <https://eim.twenkid.com/old/4/29/pred4.htm> Also lecture slides from university course in 2010:

https://research.twenkid.com/agi/2010/en/Todor_Arnaudov_Theory_of_Hierarchical_Universal_Simulators_of_universes_Eng_MTR_3.pdf

ChatGPT said:

Certainly — I can include the material from Todor Arnaudov's broader "Theory of Universe and Mind" (TOUM) framework — with its multi-scale, pancomputational/panpsychist intuition — and show how it interacts with both his 2004 "Analysis of the meaning of a sentence..." work and the 2024 paper by Johannes Jaeger et al. So the comparison will now consist of **three** pieces:

- Arnaudov's TOUM (2001-2004) — the foundational meta-theory of Universe/Mind, computation, multi-scale simulators.
- Arnaudov's 2004 "Analysis of the meaning of a sentence..." (Parts 1-4) — application of the meta-theory to semantics, machines, motivation, reinforcement.
- Jaeger et al. (2024) "Naturalising relevance realisation..." — current theoretical work on relevance, non-computational agency, cognition.

I'll highlight how TOUM's multi-scale / pancomputational / panpsychist ideas underpin Arnaudov's 2004 work, and then how that contrasts or complements Jaeger et al.

1. The TOUM meta-theory: Universe & Mind

What TOUM claims

- Arnaudov proposes that the Universe (and Mind) can be seen as a hierarchical system of "universal simulators of virtual universes" (or "sub-universes") — each level has causality-control units (CCUs), prediction/causation modules, layered levels of representation. [GitHub+1](#)
- He adopts a pancomputational or "the universe is a computer" view — but not in the simple Turing-machine sense only, rather: *all levels of reality (physical, biological, cognitive, artificial) implement computation, simulation and prediction but at different scales, precisions, domains.* [GitHub](#)
- Likewise there is a pan-psychist or "mind everywhere / scale-free cognition" flavour: even lower-level systems (cells, sub-systems) are CCUs with prediction/actuation functions; thus mind/intelligence is not only in the human brain but in many hierarchical levels. [GitHub+1](#)

- Multi-scale: The theory emphasises multi-range (temporal, spatial), multi-precision, multi-modality predictions. Intelligence/Mind is defined as “a system performing multi-scale, multi-range, multi-precision, multi-modality, multi-domain ... hierarchical prediction-causation of the future”. [GitHub+1](#)
- The “Universe Computer” metaphor: The universe itself computes/simulates sub-universes, agents and meanings; machines (thinking machines) are special cases of this modelling. [GitHub](#)
- Also there is emphasis on *matching, mapping, simulation, virtual universes* inside universes, and nested agents/selves. [GitHub](#)

Relevance to semantics, meaning, machines

- Because of the above, Arnaudov’s later 2004 Semantic/AI work assumes machines (thinking machines) are just one level of simulator; meaning for them arises via local stimuli, prediction/cause/control units, embedded environment.
- The 2004 work thus inherits TOUM’s multi-scale modelling: e.g., a thinking machine has a knowledge base, it interacts with stimuli, it builds predictions, it chooses actions (reinforcement, reward). The same conceptual machinery applies also to humans and biology and universe levels.
- The theory gives Arnaudov’s machine semantics a “cosmic” backdrop: meaning, intelligence, agency are not only in humans but in universal cosmological/cybernetic processes.

2. Arnaudov’s 2004 “Analysis of the meaning of a sentence...”

How this builds on TOUM

- In Part 1 he uses semantic analysis of a sentence in a machine context: the machine must interpret “time flies” — search the local environment, match stimuli → meaning. This illustrates the CCU/agent embedded in environment motif from TOUM.
- In Parts 2-4 he moves to reinforcement, motivation, cumulative reward, hierarchical control units (higher/lower). These align with the multi-scale control units in TOUM: lower units handle local stimuli, higher units plan for long horizon, switch domains, select goals.
- He explicitly states the importance of “local and specific stimuli, not general ones” (Part 3). This resonates with multi-scale and multi-range: local stimuli are low-level CCU tasks; general stimuli might be higher-level units.
- Intelligence is defined (in Part 4) as “search for the biggest cumulative reward for a given period ahead, based on given model of the rewards”. This aligns with TOUM’s prediction/causation hierarchy: a CCU predicts future, chooses action to maximise reward in its sub-universe/time-range.

Thus the 2004 work is grounded in TOUM: machines = CCUs in the “Universe Computer”; semantics arises via embedded interaction, local stimuli, hierarchical control, multi-scale prediction.

3. Jaeger et al. (2024) “Naturalising relevance realisation...”

Core thesis

- They argue that living agents (organisms) perform *relevance-realisation* in open, ambiguous “large worlds” (versus algorithmic “small worlds”). They claim this process cannot be fully captured by algorithmic, computational models.
 - They propose an embodied, ecological, emergent framework: autopoiesis, anticipation, adaptation, agent-arena dynamics.
 - They position cognition, agency, meaning-making as fundamentally embedded and emergent, not reducible to traditional computation.
-

4. Comparison across the three (TOUM → 2004 Arnaudov → Jaeger)

Here’s how they align and diverge, especially focusing on the multi-scale / pancomputational / panpsychist aspects.

Feature	TOUM (Arnaudov foundational)	Arnaudov 2004 Semantic/AI work	Jaeger et al. (2024)
Multi-scale hierarchical modelling	Central: CCUs at many levels, nested simulators, multi-range predictions.	Explicit: machines with hierarchical control units (higher/lower), time-horizons, local vs global stimuli.	Implicit: living agents embed multiple scales (organism, environment), but less formally hierarchical control units described.
Pancomputational / Universe as computer	Explicit: universe computes sub-universes, agents are simulators.	Machines (thinking machines) assumed computational/sub-universe.	More sceptical of computation: emphasises non-algorithmic emergent processes, argues cognition not fully computational.
Panpsychist / mind everywhere	Yes: cognitive-control units at many scales, not only humans.	Machines + biological agents both treated as CCUs; meaning & motivation apply to both.	Focuses on living agency rather than all systems; downplays machine/computation as full explanation.
Local vs general stimuli, embeddedness	Embeddedness and environment important; local stimuli matter in lower CCUs.	Explicit: “motivation is dependent on local and specific stimuli, not general ones”.	Emphasises embodied agent/arena; local ecological affordances matter.
Reinforcement / reward / goal -state model	Prediction, causation, control; reward/goals implicit in CCUs.	Explicit: reward/cumulative reward optimisation model for intelligence.	Goals/affordances considered, but not formalised purely as reward optimisation; emphasises emergent meaning and relevance, not just reward.
Role of semantics/meaning	Meaning arises via matching, simulation	Analysed: semantic analysis of sentence via	Meaning arises via relevance-realisation,

Feature	TOUM (Arnaudov foundational)	Arnaudov 2004 Semantic/AI work	Jaeger et al. (2024)
Computation/ algorithmic limits	across levels, CCUs interpret their sub-universes.	machine knowledge base + environment search.	agent frames world; less focus on machine semantics.
	Computation is broad (universe as computer), but emphasises simulation/prediction at many levels; less explicit about limits of algorithmic vs non-algorithmic.	Machine model is algorithmic/operational : knowledge base + search + reinforcement.	Strong claim: algorithmic models cannot fully capture relevance realisation; non-computational aspects exist.

5. Addressing specifically the multi-scale nature and how that interacts with the 2024 paper

- Arnaudov’s multi-scale model (TOUM) proposes that intelligence/Mind emerges from hierarchical prediction/causation at many levels: e.g., cells, organs, brain, society, universe. This means that semantic analysis of a sentence (Part 1) is just one level of operation; motivation/reinforcement (Parts 2-4) map across levels (local stimuli → higher goals).
 - In the 2004 paper he uses that multi-scale logic: for example, a machine (thinking machine) has local stimuli and higher goals over time horizon; the environment is a simulator/sub-universe; there may be nested control units. Thus the “banker buying a donut when rich” example (if present) could be mapping between higher-level goals (wealth maximisation) and local stimuli (street donut).
 - For Jaeger et al., one could say: they emphasise that relevance realisation necessarily spans levels (organism, environment, ecological niche, cultural/historical embedding) and hence multi-scale. However, they emphasise *that computational, algorithmic approaches* (which often assume fixed scale/time horizon, formalised worlds) cannot handle the openness of “large worlds” and the emergent linking across scales. Thus their critique can be seen as a challenge to models like Arnaudov’s (which are algorithmic) to scale to truly multi-scale, open-ended embodied agents.
 - In particular: Arnaudov’s view would say: yes, a thinking machine (or human) runs nested prediction modules at many scales; but Jaeger et al. would ask: can an algorithmic machine truly realise relevance across all scales (from micro-stimuli to macro-contexts) in the way living agents do? They lean toward “no, at least not fully”.
 - So the interplay: TOUM and Arnaudov’s 2004 work propose a broad multi-scale computational/agentic model; Jaeger et al. propose that for living cognition, something goes beyond that computational multi-scale model (which would include machines) — i.e., algorithmic methods may not suffice.
-

6. My refined assessment

- My earlier summary under-emphasised the richness of Arnaudov's multi-scale, local stimuli, hierarchical control model. You are correct that those themes are present (especially local stimuli vs general ones) and important.
- What still remains is that Arnaudov frames his solution within an operational/computational scheme (knowledge base, machine, reinforcement, reward) *inherited* from TOUM's universe-as-computer view. By contrast, Jaeger et al. emphasise the **limits** of such algorithmic/computational models when applied to real biological cognition in open worlds.
- Thus Arnaudov offers a constructive model (for machines or humans) of meaning, motivation, intelligence across scales; Jaeger et al. offer a *critical* perspective, arguing that some aspects of cognition (relevance realisation) may escape purely algorithmic frameworks.
- The multi-scale nature is central to both — but Arnaudov uses it to build machine/agent models; Jaeger uses it to show complexity and emergent features that challenge classical computational models.

If you like, I can prepare a **detailed mapping table** with **specific quotes** from TOUM (2003/2004 treatises), Arnaudov 2004 (Parts 1-4) and Jaeger 2024, highlighting how each deals with e.g. multi-scale prediction, local stimuli, relevance, computation vs algorithm, meaning. Would that be helpful?