Proceedings of Machine Learning Research 131 (2020) 32-61 IWSSL

Seed-Programmed Autonomous General Learning

Kristinn R. Thórisson THORISSON@RU.IS
Department of Computer Science, Reykjavik University & Icelandic Institute for Intelligent Machines Reykjavik, Iceland

Editors: Minsky, H. and Robertson, P. and Georgeon, O. L. and Minsky, M. and Shaoul, C.

Abstract

The knowledge that a natural learner creates of any new situation will initially not only be *partial* but very likely be partially *incorrect*. To improve incomplete and incorrect knowledge with increased experience – accumulated evidence – learning processes must bring already-acquired knowledge towards making sense of new situations. For the initial creation of knowledge, and its subsequent usage, expansion, modification, unification, and deletion, knowledge construction mechanisms must be self-guided, capable of self-supervised "surgical" operation on existing knowledge, involving among other things self-inspection or *reflection*. Further, the information that makes up an agent's knowledge set must thus be structured in a way that supports reflective processes including discrimination, comparison, and manipulation of *arbitrary subsets of the knowledge set*. Few proposals for how to achieve this in a parsimonious way exist. Here we present a theory of how systems with these properties may work, and how cumulative self-supervised learning mechanisms can reach levels of autonomy like those seen in individuals of many animal species. Our theory rests on the hypotheses that learning is (a) organized around causal relations, (b) bootstrapped from observed correlations, using (c) fine-grain relational models, manipulated by (d) micro-ampliative reasoning processes. We further hypothesize that a machine properly constructed in this way will be (e) capable of *seed-programmed autonomous generality:* The ability to apply learning to any phenomenon – that is, being domain-independent – provided that (f) the seed reference observable variables at "birth", and that (g) new phenomena and existing knowledge overlap on one or more observables or inferred features. The theory is based on implemented systems that have produced notable results in the direction of increased general machine intelligence.

Keywords: Cumulative Learning, Self-Supervised Learning, Understanding, Seed Programming, Machine Learning, Knowledge Representation, Autonomous Generality

1. Introduction

Many examples of self-supervised autonomous learning are found in nature—in fact, a notable part of adaptive behaviors in individual organism with a brain is acquired through experience, over varying periods of time. Of special interest to artificial intelligence (AI) researchers are methods that would allow a machine to amass such adaptations, as well as generalize these over time, to be applied as appropriate in future circumstances.

Common to all cognitive learners in nature is that they started from a seed and grew into thinking beings. Out of the path that brings a single cell to a fully grown animal, we address here the subset of how a newborn learner may bootstrap its knowledge in the environment it is born. What allows a learner to autonomously generate knowledge of itself and its environment from a kind of cognitive seed—an initial and necessarily small "knowledge nugget"? How does it hold on to what it learns over time, in a way that future learning can be built on? And how does it modify it, in light of new information? Such *autonomous cumulative learning* is markedly unlike most machine learning (ML) methods developed to date and not well understood at present (Thórisson et al. 2019).

In this paper present a theory of how self-supervised learning of novel phenomena may be realized. We consider the conjecture that knowledge bootstrapping at birth is a special case of the general principles involved in bootstrapping learning in partially-unknown, *novel*, circumstances. In both cases a learner starts with something given that is inadequate and insufficient for addressing the novelty, facing the cognitive task of making use of what it already has, to make sense of it (Wang 2013). Addressing this subject will require considering three interlocked realms: (a) The world of the learning agent and its target *task-environments*; (b) the mechanisms for control and management of the *cumulative learning* process; and (c) how learning is bootstrapped through existing information—a *knowledge seed*.

The mechanisms of learning found in nature[1] are the result of requirements imposed by a non-axiomatic and combinatorial nature of the physical world itself: Its underlying rules can only be inferred from observations of the transformations of measurable variables accessed through sensory organs. Yet such learning happens in a self-supervised manner—without the help from a parent, teacher, paper-and-pencil, pocket calculators, or other outside assistance. Individuals of many animal species display robust autonomous properties that could benefit artificial learners, including learning many tasks over time, transferring knowledge between situations (transfer learning), handling distractions, learning many tasks (multi-goal learning), and handling novelty. Yet most of these remain nonexistent in modern machines.

At its simplest, learning a single well-defined task incrementally might be accomplished in a machine using methods and steps provided beforehand by a human coder. Should some part of a task or environment be unknown beforehand, however, autonomous methods for incremental knowledge unification would be needed. An important part of cumulative learning is relating relevant new information to what already exists, as for instance when the features of a particular outdoor sport, say tennis, are related to a different sport, e.g. football, with which it shares some similarities, allowing a quicker and more coherent knowledge creation and subsequent understanding of the new phenomenon. A large part of such processes must involve analogies (cf. Sheikhlar et al. 2020, Besold and Schmid 2016) – not necessarily the kinds that we make consciously when reading poetry or playing compare-and-contrast games, but rather more automatic – such that the relevant features and similarities to be identified and analyzed are automatically highlighted. As we shall

1. To keep the discussion focused, we consider humans a key example of what we would our artificial agents to reach in terms of cognitive abilities, but some household and even wild animals have demonstrated abilities worthy of consideration as well (cf. Balakhonov and Rose 2017, Patterson and Gordon 2001) and are perhaps more appropriate for comparison and contrast at the present stage of AI development.

argue, in addition to analogies the process also involves three other forms of reasoning, deduction, abduction and induction.[2]

The more *general* a cumulative learner is, the more diverse tasks and environments it should be capable of learning, other things being equal, and the more diverse its acquired knowledge will become over time. Such learning, if appropriately implemented (a focus of this paper), should make an agent increasingly capable of handling a growing number of unfamiliar situations—directly, without blind experimentation. For that to be possible, a continually active process of assessing the relevance of acquired knowledge for any and every situation must be efficient and effective.

While some of the topics mentioned, e.g. generality and autonomy, are already in the cross-hairs of some research programs (cf. Reid et al. 2018, Lawless et al. 2017, Besold and Schmid 2016), relatively few aim explicitly to create a single learner that captures them all. Yet that is what is required to realize general machine intelligence. One reason for lack of progress on this point may be the perceived difficulty of addressing the large set of requirements at once that such systems would need to meet (cf. Thórisson et al. 2015, Laird and Wray III 2010). The application of a modular approach seems at first glance to be in accordance with the scientific method, but for any complex system with many parts, it is not sufficient to study the parts in isolation. As anyone familiar with the internal combustion engine, in a system that relies on a large set of interlocked, intricate mechanisms that produce a coordinated output, the mechanisms must be tuned to each other for the whole system to work—even a tiny deviation (sand in a ball bearing; a loose electrical connection; a splash of water in the gas tank) can cause a breakdown. This principle is equally important in reverse-engineering more complex phenomena such as thinking. Dissecting the mind into small pieces and studying them separately may produce some results, but insights into how the system as a whole works resulting from such an approach will be impossible, forever prevented by incongruent isolated theories, each resting on background assumptions orthogonal to the others in one or more ways—any hope of ever piecing together a pot-pourri of parts developed this way into a coherent architecture is wishful thinking (cf. Thórisson 2012, Thórisson 2008, Garlan et al. 1995).

Answering Newell's (1994) call for unified theories of cognition, our theory outlines how all these properties may be unified in an effort to realize artificial agents with general intelligence. The theory follows a *constructivist-AI methodology* approach (CAIM) outlined in Thórisson (2017, 2012) and combines several threads in my and my team's work over the past ten years on the topics of *cumulative learning* (Thórisson et al. 2019, Thórisson and Talbot 2018a, Thórisson and Talbot 2018b), *autonomous transfer learning* (Sheikhlar et al. 2020), *self-programming* and *autonomy* (Nivel et al. 2014b, Nivel et al. 2014a, Nivel and Thórisson 2013a, Nivel and Thórisson 2008), *bounded recursive self-improvement* (Nivel et al. 2013b), *understanding* (Bieger et al. 2017, Thórisson et al. 2016b, Steunebrink et al. 2016, Nivel et al. 2013b), and *task theory* (Eberding et al. 2020, Bieger and Thórisson 2018, Thórisson et al. 2016a).

Most of the concepts and principles described here have been implemented and tested in systems whose demonstrated abilities have exceeded in some aspects other learners on

2. Again, it may be useful to think of these as low-level (subconscious) processes—we use the prefix "micro" because in our theory they are fine-grain and fast and much of their operation may not be readily accessible through familiar conscious introspection methods.

dimensions including sequence learning, multi-task learning, natural language interpretation and continuous lifelong learning (cf. Nivel et al. 2014c, Nivel et al. 2013a, Hammer and Lofthouse 2020, Wang 2007), but especially with respect to realizing all of them in a unified way in a single learner. This, as well as relevant work of others, will be referenced throughout the paper. Of particular relevance is our work on the S0 and S1 agents, constructed in the AERA framework (Nivel et al. 2013c), which demonstrated how the principles described herein can enable a machine learner to learn highly complicated real-time tasks (interaction with humans) by observation. In particular, our S0 agent learned a complex goal-driven task (complying with commands provided in natural language and gesture) in only 2.5 minutes; S1 derived sufficiently detailed knowledge from observing humans in an interview about recycling various materials, also based on natural situated communication but with a much more complex 100-word vocabulary and free-form grammatical sentences. We give a short description of these agents in Section 6.

The paper is organized as follows. First we will present a high-level overview of the theory, including some background concepts, with a particular view to 'generality' and 'autonomy' (Section 2). Then we will look at what kinds of worlds the present work is relevant to (Section 3). After this we look at the concept of seed-programming (Section 4). Then we discuss cumulative learning (Section 5), which contains four subsections, each focusing on key aspects of such learning (discussing modeling, semantic modularity, causal-relational[3] models[4] and micro-ampliative reasoning, respectively). We conclude with a summary of the preceding sections (Section 6) and a relatively short conclusion (Section 7).

2. Autonomous Generality

We consider the effort to create systems with "general intelligence" really to be about implementing systems with *autonomous generality*, since the requirement to handle an increasing variety of situations, environments, and worlds grows (i.e. increase in generality), the requirement on autonomy increases too, as a generally intelligent agent that cannot think independently is by definition not general.[5] Conversely, to achieve high levels of autonomy in complex partially-observable worlds requires increasing levels of intelligence. The concept of 'generality' in the context of intelligence can be addressed from various angles, a common one being in relation to variety of one sort or another that an intelligent agent can handle—kinds of tasks, data, situations, environments, domains, or 'worlds' (Thórisson and Helgason 2012). While an absolute measure or scale of generality would nice to have, this is not very doable without a solid theory on which to rest such a scale (cf. Goertzel 2015). At the very least there should be some way to say that one AI is more general than another. This, however, may be even more complicated than would seem at first, because generality is a multi-dimensional concept.[6]

3. We define causation in a practical way, meaning the useful coupling of being able to achieve X by doing Y (cf. Pearl 2009, Pearl 2001).
4. The meaning of the term 'model' in our use is comparable to that of Wang (2006a) and cybernetics (cf. Ashby 1956). This is further discussed in Section 2.
5. Thanks to Leonard Eberding for this angle (personal communication).
6. For instance, is an agent that can learn many things but takes a very long time to do so more general than an agent that can do 1000 tasks extremely well but can hardly learn, or does learning always trump staticity?).

Assuming a capacity to *learn* as given in any system classified as intelligent, and generality to refer to an enumeration of variety, what then is autonomy? Autonomy has been associated with artificial intelligence for decades and has commonly made an appearance in robotics and multi-agent systems, where it has been analyzed both with respect to tasks and capabilities (cf. Beernaert et al. 2018) and cognitive capacity and architecture (cf. Thórisson and Helgason 2012, Wooldridge and Jennings 1995). While both practical and theoretically useful for comparing different systems, the reach of this work typically does not include the concerns of biologists, whose interest in understanding autonomy from a self-organizing auto-catalyzing perspective aims to uncover the principles of autonomy in nature, especially life itself (cf. Winning and Bechtel 2019, Letelier et al. 2011).

We may think of autonomy as spanning at least three grand levels of sophistication: First, the lowest level may be called the "mechanical" level of autonomy—here we would list the familiar Watt's Governor and old-fashioned thermostats, as well as modern deep neural networks and the like. Their function is fixed after they leave the lab, which is why some feel even hard-pressed to call them autonomous. These systems may nevertheless "perform complex tasks" unaided (before they break). The word "automation" is often used when referring to systems at this level. At the middle level we put system with some cognitive autonomy: The ability to adapt and handle novelty, to "figure things out" and even create new ideas. At this level we see humans and higher-level animals including some dogs, crows, and parrots. The highest level is the biological one, where life springs forth and spawns the other two levels. It is in our view the "most autonomus" one – should we try to pit them against each other – because it is a prerequisite for the others. However, it is a distraction to those focusing on artificial intelligence because intelligence calls for discretionarily constrained adaptation: The ability of the system to constrain its own behavior by choice, through selecting goals, sub-goals and other factors at its own discretion, through reasoning and logic (Thórisson 2020). On the border between the mechanical and cognitive levels of autonomy lie (basic) reinforcement-based learners that, while able to change their function at runtime, are limited to a handful of variables and cannot handle novelty.

In this paper we look at autonomy from the perspective of knowledge representation and acquisition at the cognitive level—especially in relation to "figuring things out." While quite possibly less challenging than life itself, it does share some of the same concerns (cf. Rocha 2010) and harkens back to the early times of cybernetics, in particular what has been called 'second-order cybernetics' (cf. Heylighen and Joslyn 2001). Some work in developmental robotics has also been focused on these subjects (cf. Smith and Gasser 2005, Weng 2004), human cognition (cf. Piaget 1953) and artificial general intelligence (cf. Wang et al. 2016, Rocha 2010, Wang 2009). Our concept of learning is fairly broad, incorporating not only the accumulation of (practical) knowledge but also the ability to use this knowledge achieve some *autonomy in learning*. Other things being equal, a system that is general and autonomous when *performing* tasks, but not autonomous in inventing or finding new solutions, will not be as generally autonomous as one can do both. If systems capable of cumulative learning can take initiative and actively uncover missing information, and even invent new concepts for the purpose of getting better, they are able to not only do things but also "figure things out." Yet if they are not sufficiently general their autonomy will be bounded to a limited set of tasks, circumstances, data formats, etc.: All of these are necessary (but not necessarily sufficient) for general intelligence.

This, then, is what we wish to create: An autonomous general cumulative learner. Our theory of autonomous general learning rests on the hypotheses that learning is

(a) organized around causal relations,

(b) bootstrapped from observed correlations, using

(c) fine-grain relational models, manipulated by

(d) micro-ampliative reasoning processes.

We further hypothesize that

(e) a machine properly constructed in this fashion will be capable of *seed-programmed autonomous generality:* The capacity to apply learning to any phenomenon without help – i.e. being autonomously domain-independent – provided that

(f) its seed references observable variables (at "birth") and that

(g) existing knowledge shares one or more observable variables, patterns or inferred features with novel phenomena to be learned.

Of course, this begs the question what is meant here by "properly," something that we aim to address in this paper. It should be noted that our use of the term 'model' does neither imply mathematical (axiomatic) models, nor any kind of logic based on model-theoretic semantics (Taski 1944, Putnam 1981); rather, like Wang's experience-grounded semantics (EGS; Wang 2006a) our models model the experience of an agent, in particular indications of causal relationships that may be found there. Unlike 'models' in artificial neural networks, from which they are also quite different, our models have internal semantics that carry down to their smallest components: patterns, made up of variables, transformation functions and relations.

To explain the background of our theory, three high-level things must be considered, as already mentioned: The world of the learning agent and its target *task-environments*; the mechanisms for control and management of the *cumulative learning* process; and how learning is bootstrapped through the information it is born with—its *seed*.

Task-Environment. As argued by Steunebrink et al. 2016, no general intelligence in a complex environment such as the physical world can be granted access to a full set of axioms of the world, let alone the $\langle system, environment \rangle$ tuple, and thus the behavior of a practical generally intelligent artificial agent as a whole simply cannot be captured formally. Our ultimate target is the physical world, which follows (local and global) rules but will always contain vastly more unobservable, unaccessible and unmanipulatable variables than those any learner – hypothetical or practical – will ever know. In such a world novelty abounds—most things that a learner encounters will contain some form of novelty. We look at this in Section 3, *Kinds of Worlds*.

Representation. In a novelty-rich world, no challenge, problem or phenomenon is identical from moment to moment, if only for the simple reason that the progression of time moves things continually from the present to the past, changing the context: Time is a semantic property (Lee 2009). A cumulative learner figures things out because it is equipped with corrigible – non-axiomatic (Wang 2006a), defeasible (Pollock 2010) – knowledge, and as it encounters novel problems it adds to its current knowledge as well as modifies it to adapt the new information. Ashby's Requisite Variety Theorem (1958) states that the representation of a controller must have a granularity as small as the smallest feature that we want the learner to discern. Our approach uses a two-part representation scheme, consisting of causal-relational models (CRMs) and patterns—both of which are information structures amenable to manipulation, comparison, and compositionality. Equally importantly, they are abstract and can thus be used a any level of detail; their capacity for abstraction comes from their ability form hierarchical sets. Key management mechanisms over this representation are (corrigible) abduction and (corrigible) deduction.

Seed. A learner that is born knowing nothing cannot learn anything, because there is nothing to tell it how (or what) to learn, the equivalent of the proverbial baby that's thrown into the deep end of the pool: It will not learn to swim if it doesn't even know how to move its arms. That natural intelligence is a blank slate at birth – a *tabula rasa* (Aristotle 350 BCE) – clearly cannot be correct. A seed must thus include *some assumptions* about the world the agent is born into, to bootstrap its learning—an "inductive bias" (Dubitzky et al. 2013) or *bootstrap program* that tells the newborn what to initially pay attention to and how to turn experience into knowledge.

In prior work we have demonstrated an implemented architecture, the Autocatalytic Endogenous Reflective Architecture (AERA, Nivel et al. 2013c), that incorporates these ideas. AERA can learn autonomously many things in parallel, at multiple time scales. Results of detailed evaluation show that AERA can learn complex multi-dimensional tasks from observation, while provided only with a small ontology, a few drives (high-level goals), and a few initial models, from which it can autonomously boot-strap its own development. This is initial evidence that our constructivist methodology is away for escaping the constraints of current computer science and engineering methodologies. Human dialogue is an excellent example of the kinds of complex tasks current systems are incapable of handling autonomously. The fact that no difference of any importance can be seen in the performance between AERA and the humans in simulated face-to-face interview is an indication that the resulting architecture holds significant potential for further advances.

3. Kinds of Worlds

What kind of world is our physical reality? Without regularity, no patterns would repeat, no event would reliably result in something specific, no categories of objects or events would exist, and no action would lead to any predictable outcome. Such a world would consist entirely of noise and no learning could take place. On the other hand, completely deterministic, fully observable worlds call for little guesswork. Whether fully deterministic or not, regularity means that some things go together more than others.[7] A world where some things reliably precede some things and not others is a *causal* world.

7. Whether the physical world is fundamentally deterministic or not "at its core" is a metaphysical question and immaterial for these purposes—what matters here is that any agent aiming to survive in it must make do with the limited information it has, as stated by the assumption of limited knowledge and resources (AIKR, Wang 2006a).

The scale of physical reality that human intelligence operates at – the *object* scale[8] – a lot of regularity is found, allowing us to survive in complex environments. An intelligent agent in the physical world must deal with (1) partial observability, (2) an enormously large number of building blocks, from glare on water ponds to the faces of relatives to sunsets by the beach, that (3) can combine in an exponentially greater number of ways over the lifetime of a learner. If we define 'novelty' as any pattern that may diverge in perceptually noticeable ways from patterns seen before, we can safely say that novelty is far more common than precise repetition (and according to (Wang 2019) is a defining property of intelligence).

From a learning agent's perspective, whose memory can only fit a minuscule fraction of the world's combinatorics in their memory (even if they were strictly limited to the object scale), any method or "trick" that makes its tasks easier is welcome. Three critically important features make intelligence in the physical world a practical possibility: Multi-level regularities, a.k.a. repeated patterns, and repeatable transition functions, a.k.a. "laws of physics."[9] The third is the proverbial arrow of time, composing what we call "reality" out of temporally morphing patterns that follow lawful transition functions. While the ocean may never repeat the same waves exactly the same way twice, at a higher level of analysis its waves share sufficient similarities to be meaningfully compared (and contrasted with other patterns in other contexts). For data measured at the lowest level of sensors (e.g. a 2D grid such as the eye's retina or a digital camera CCD), this hierarchy of patterns makes what otherwise would present too much novelty to keep track of in any practical manner analyzable at higher levels in groups or blobs, dissecting a visual scene into trees, mountains, buildings, and people, each of which in turn can be seen as consisting of parts such as branches, roofs, ears, noses, etc. Similarities between levels forms yet another dimension of similarity comparisons, presenting in essence a fractal dimension—yet another "trick" that intelligence can use to handle complexity. As has been pointed out before (Richardson 1998), the multitude of factors that matter to an autonomous agent in the physical world can be thought of as forming a *dynamic hypergraph*, where values of sets of particular variables dictate the values or value ranges of others (one need only think of the way that walls of a room limit the possible positions of objects within the room).

The world W consists of a set of variables $\mathcal{V}$, dynamics functions $\mathcal{F}$, an initial state $\mathcal{S}_0$ and relations $\Re$ between these: $W = \langle \mathcal{V}, \mathcal{F}, \mathcal{S}_0, \Re \rangle$. The variables $\mathcal{V} = \left\{ v_1, v_2, \ldots, v_{\|\mathcal{V}\|} \right\}$ represent all the things that may change or hold a particular value in the world. The dynamics can intuitively be thought of as the world's "laws of nature," continually transforming the world's current state into the next: $\mathcal{S}_{t+\delta} = \mathcal{F}(\mathcal{S}_t)$. The concept of 'domains' as subsets of the world, $\mathcal{D} \subset W$, where a particular bias of distributions of variables, values and ranges exists, may be useful in the context of tasks that can be systematically impacted by such a bias (gravity versus zero-gravity). Each variable v may take on any value from the associated domain $d_v \in \mathcal{D}$; for physical domains we can take the domain of variables to be a subset of real numbers.

8. The spatial lower bound of (unaided) sensation/perception is not much under 1 millimeter; the temporal lower bound is around 10 milliseconds. The lower spatial (unaided) resolution of the perception-action loop is a few millimeters and temporally around 100 msecs. If we limit the upper bound to our maximum lifespan all those dimensions would max out at 100 years. The vast majority of human cognition and perception-action events span the range from 1 to 10^4 seconds.
9. Calling them "laws" is actually misleading—to be certain of their status as *laws* would require a third-person view of the universe and our existence, something we can of course never attain.

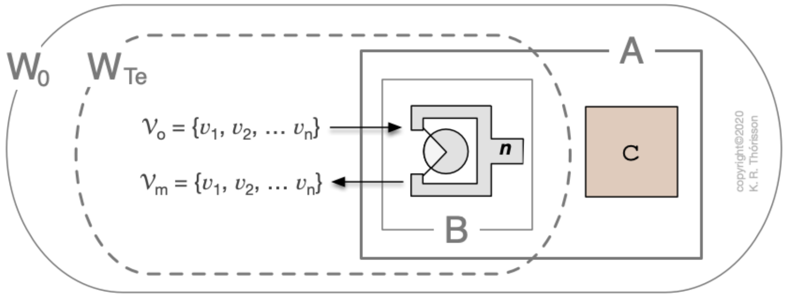

Figure 1: An agent A in a task-environment embedded in a world ($Te \subset W_0$) consists of a controller c and a body B consisting of one or more (n) transducers that can observe accessible variables ($\mathcal{V}_o$) and affect manipulatable variables ($\mathcal{V}_m$). While the body belongs to Te we may think of the controller to be either outside of it or part of it (the latter will be true of any physically implemented controller).

A phenomenon in the world is any grouping of variables and relations in the world that we choose to group as such; $\Phi = \langle \mathcal{V}_\Phi, \Re_\Phi | \mathcal{V}_\Phi \subseteq \mathcal{V}_\mathcal{W} \wedge \Re_\Phi \subseteq \Re_\mathcal{W} \rangle$. It consists of elements $\{\varphi_1 \dots \varphi_{\|\Phi\|} \in \Phi\}$, which can themselves consist of other phenomena, variables, and relations $\Re_\Phi$ (causal, mereological, etc.). $\Re_\Phi$ couples elements of Φ with each other, and with those of other phenomena in the world (Bieger and Thórisson 2017, Thórisson et al. 2016b), and can be partitioned into inward facing relations $\Re_\Phi^{in}$ between element pairs $\varphi_i, \varphi_j \in \Phi$ and outward facing relations $\Re_\Phi^{out}$ between element pairs $\varphi_i \in \Phi$ and $\beta_j \in W$.

An environment e is a view or perspective on the world, $e \subset W$; a *task-environment* Te is an environment where a set of tasks can be performed and where some variables are observable ($\mathcal{V}_o$) and manipulatable ($\mathcal{V}_m$)—parts of the world which may change dynamically as the state-space of the hypergraph it forms is traversed, based on relations between variables. The state-values of any element or variable in Te at any point in time is determined by a causal chain that we can think of as a directed acyclic graph. Any complex environment is one in which the total number of variables is vastly greater than the total number of observable variables $\|\mathcal{V}\| \gg \|V_o\|$ and observables more numerous than manipulatables $\|\mathcal{V}_o\| > \|\mathcal{V}_m\|$, since these factors directly contribute to a system's complexity from an agent's limited view.

3.1 Agents in Task-Environments

We define an *embodied cognitive system* as consisting of two main interacting components: A *goal-driven learning controller* of a body and a task-environment.[10] By "body" we mean constraints that limit what the agent can do and how, and what it can sense (measure) and how, at any point in time, through transducers (sensors, actuators) that interface with the world. By "goal" we mean a set of target variables and associated values, with permissible deviations and time limits, as well as potential constraints ("negative" goals: what must be avoided while working on the task). By "goal-driven" we mean that the system actively

10. This formulation is commonly used in control theory and reinforcement learning, but is general enough to also cover more "traditional" cases of supervised and unsupervised learning.

seeks state spaces that meet certain conditions, as specified by its goals; by "learning" we mean that experience may be used to direct subsequent behavior in favor of such goal seeking, to do it better (e.g. fail less often upon repeated attempts, or get closer to the target values of the goal with practice).

A basic task T in a particular environment Te is any specification that consists, at a minimum, of an initial starting state $\mathcal{S}_0 = \{v_1, v_2, ...v_n\}$ (variables in the environment having particular values) and time t_{go}, a top-level goal $G_{top} = \{S_1 S_2, ...S_n\}$ (a set of states with desirable joint variable assignments that must be achieved at particular times) and sub-goals, $\mathcal{G}_{sub} = \{g_1, g_2, ...g_n\}$ that serve to further specify G_{top}, constraints $\mathcal{G}^-$ (i.e. negative goals—states that result in instant failure and should be avoided), a body $B = \{b_1, b_2, ...b_n\}$ (an interface to W including a set of transducers with a reasonably stable relationship and accompanying affordances[11] and co-located with the controller), an end time, t_{stop}, and possibly some additional information about the task, e.g. instructions, I: $T = \langle \mathcal{S}_0, \mathcal{G}_{top}, \mathcal{G}_{sub}, G^-, B, t_{go}, t_{stop}, I \rangle$. An *assigned* task will have all its variables bound and reference to an agent that is to perform it. Complex tasks may be thought of as being composed of sub-tasks, $T_1 = \{T_a, T_b, ...T_n\}$, forming a web of relations through one or more shared variables.

When performing a task an agent may sample values of variables from the environment, produced by its transfer functions $\mathcal{S}_{t+\delta} = \mathcal{F}(\mathcal{S}_t)$, through measurements via its transducers; we say that a *percept* p is a subset of variable values, $p = v \subset \mathcal{V}$, that are measured by an agent's transducers at some point in time (or over some period) and accessed by its controller (i.e. perceived), and that $\mathcal{P}_{\varphi(t)} = \{p_{1(t)}, p_{2(t)}, \ldots p_{n(t)}\}$ is a set of percepts related to one or more elements $\varphi \in \Phi$ at time t. For the kinds of task-environments are interested in, the number of exposable variables at any point in time is typically vastly larger than observable ones, $\|\mathcal{V}_{e(t)}\| \gg \|\mathcal{V}_{o(t)}\|$. Only variables that are *exposable* can be observed, and only variables observable at time t can be part of a percept; thus, observability is a temporal property of variables but exposability is not. Exposable variables may be unobservable for various reasons and at various times (most commonly because they are obscured by objects or because they simply are not in view of the agent's transducers, but also if they require a particular technology to be invented to be seen, e.g. a telescope or microscope.) We may assume that percepts are always in the "now"; a "percept" that is not being delivered directly by a sensor modality is not a percept but rather an encoding or representation – a memory – of a percept.

Hidden variables may become visible and others may become hidden as an agent interacts with phenomena in the world, exposing their relationship through their transition functions via systematic state comparisons of time periods. Any phenomenon that a learner may encounter and of which it has no knowledge is by definition novel to the agent, Φ_{nov}. As an agent learns, that novelty turns into familiar information, Φ_{fam}, and elements of Φ cease to be novel, $\Phi_{fam} = \Phi - \Phi_{nov}$, as knowledge grows over time. We will come back to this shortly. In the kinds of task-environments we are interested in, for any agent, the set of novelty is vastly larger than that of what is familiar, $\|\Phi_{nov}\| \gg \|\Phi_{fam}\|$, even after lifelong learning. It is therefore given, in the case of the agenthood we are interested in, that the set of unknown phenomena, at all times, is huge.

11. "Affordances" (Gibson 1966), in our conceptualization, are quite simply *familiar verified causal chains*. In the case of transducers this involves the primitive that can flow in and out of them (input and output).

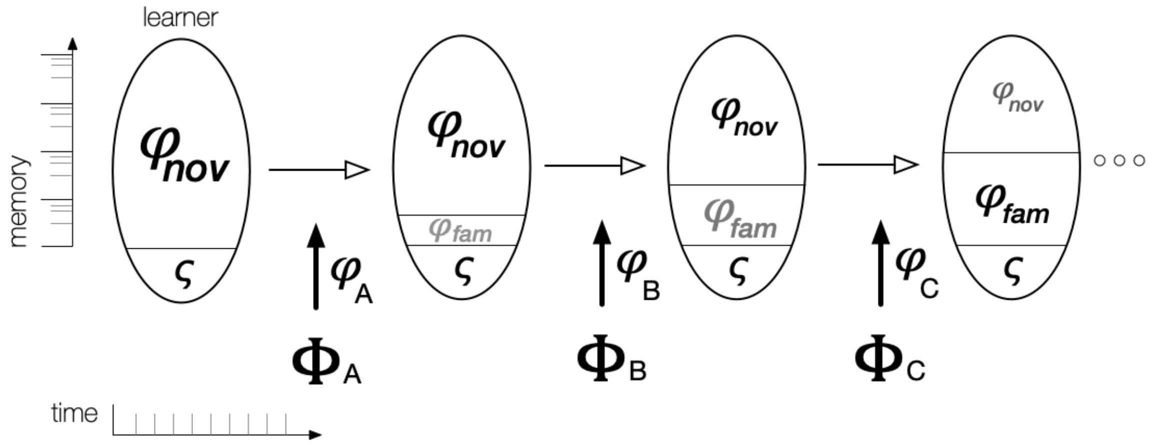

Figure 2: The amount of knowledge a learner can hold (represented by an oval) is bounded. At the outset the only knowledge is contained in a seed (ς). As a learner encounters novel phenomena over time, $\Phi_A, \Phi_B, ...\Phi_n$, it creates models of what it observes (from its percepts of them), bootstrapped at "birth" (leftmost oval) by the seed (ς), and later by newer knowledge, Φ_{fam}. Over time, the amount of its percepts related to aspects of phenomena it can predict (Φ_{fam}) increases; reciprocally, the amount of phenomena it encounters that are novel (Φ_{nov}) decreases. The seed (ς) does not change; its models will either continue to be used or be superseded by newer knowledge. As knowledge builds up the knowledge management mechanisms must increasingly rely on compaction (compression) methods to fit more knowledge into the limited memory (cf. Thórisson et al. 2019). This will involve some forms of induction and forgetting, among other methods.

4. The Need for a Seed: Seed-Programmed Autonomy

Human knowledge of the natural world – indeed, that of any learner in the physical world – is the result of a composition process, pieced together incrementally from experience with the world over time, accumulating in a somewhat systematic way. This is the way any learner must operate where complete information is not available up front, at birth. If an agent is to learn independently, autonomously, its knowledge acquisition processes must be self-guided—it must have *existential autonomy*. To successfully engineer artificial systems with this capacity requires uncovering principles of such *cumulative learning*.

If we wish to create an autonomous learner that can operate without human intervention, or some other kind of outside interference, reprogramming etc., such a system will by definition have to be "born" with a set of initial knowledge—a *seed* from which the learner's knowledge should grow cumulatively, as it interacts with the world, in a self-supervised fashion. If we define a "newborn" cognitive agent as one with only a seed and no experience of interacting with the world, a newborn's first task is to bootstrap its own knowledge acquisition process, starting from what's in its seed. What kind of information must be in such a seed? If we are designing a system from scratch it will be up to us to decide; however, if the aim is to make the system autonomous the information must be sufficient for the system to bootstrap its knowledge, our hands are tied after we bring it into the world, so we must choose wisely...

The purpose of intelligence is to get stuff done. For natural intelligences this involves surviving and making offspring (lest the species die out); for an artificial agent the main purpose – we call this top-level goal or "drive"[12] – may be to manufacture computer parts, design new manufacturing plants, or invent new ways to compute. Intelligence is a practical solution to practical problems, and few things are as practical as *taking action*,[13] to get something done, preventing something from happening or affecting the world in some way—through deliberate achievement of *goals*. If intelligence is not good for taking some sort of action in a time of need, what is it good for? (Absolutely nothing.)

What kind of action should a seed enable a newborn to take? A learning agent's motivation for acquiring and improving its own knowledge – its impetus for modeling the world – must exist at birth, otherwise no learning will take place. To work, such motivation must reference actionables and observables in the world. This leads to the following principle:

§1 *The seed of an autonomous general learner must include at least one (top-level) drive that references variables necessary for grounding.*

For an agent constructed to do more than one tasks – let's just say various household chores, for the sake of argument – could we not provide it with a detailed description of every such possible task up front—dusting, vacuum cleaning, loading and unloading the washing machine, etc.? If we had to do that its seed would have to contain an impractical amount of detail (think of the all the ways any even just a tiny portion of a lacquered floor may look, from different angles, under different lighting conditions, throughout the year). That is not the only reason though for this being impractical; such a seed – if we could pull it off – would be very sensitive to minor changes in any task and environmental variables; anything that is slightly different from the seed would stop the agent from being able to achieve it successfully. This means the seed must contain generalizations in the form of rules. (We will discuss the form they may take shortly.) The main reason for populating the seed with general rules is that not everything may be known up front—even the set of tasks that we might want the agent to perform may change drastically in the future. In a complex world with infinite combinatorics, this is eventually bound to happen: Novelty is guaranteed and thus must be assumed when constructing a seed.

We have uncovered a general and fundamental principle of general intelligence: A general learner must be able to address novelty. If the seed is to be of use in bootstrapping a *general learner* it must appeal to some *general principles* of the world into which the agent is born. The more general we want an agent to be, the more general the initial rules at birth must be—remember, at birth the seed contains *the totality of the agent's knowledge*. The more specific the rules are, the more sensitive they are to future changes in conditions, tasks, and environments. Could we start the general learner with an empty seed? No, because the purpose for intelligence is to control action, and the learning that the seed is supposed to bootstrap must eventually lead to a *discovery and verification of ways of affecting the world*. A seed must thus contain some instructions precisely about that – affecting the world – in particular, how *this particular agent* can affect the world via *its own body*, which is its interface to the world (physical forces). The seed code must reference parameters that are assumed to be actionable in the learning domain, through manipulatable variables

12. It deserves its own name because in a general learner it serves a different role than a typical task goal or subgoal: It ensures the learner's *purpose*, and is likely to be implemented in a somewhat different, possibly more rigid way.

13. In a world with a real-time clock, inaction also has consequences and may be used to get things done—or mangling them.

(e.g. outer limbs and movable sensors). This is necessary but not sufficient: To know whether one action is good and another one bad, the seed must be populated with some general reference to observable variables—measurable entities whose existence in the birth world is fairly certain; variables that can affect the agent's sensory aparati. Lastly, the seed must supply some ways in which actions and observations may be associated or related in an information-crowded world.

In the physical world, the most fundamental principles known are the laws of physics, as they have slowly been uncovered in fits and spurts over the past 2000 years. One problem with the laws of physics, even though many of them may be represented compactly, is that they are very general. In science that's a feature of course, not a bug, but in the context of a seed it means that they do not specifically reference anything that the agent can directly sense or act on (that's in part why it took so long to uncover them). Humans were not born with knowledge of these—they had to discover/invent them. A learning agent in the physical world will not only be learning sub-divisions of what it sees and hears, it will be learning about how things change over time. Providing rules about relations between action and perception is difficult without using complex concepts—concepts that most certainly will be beyond a newborn's ability to grasp. The agent learns about the world little by little, over time, through its experience, but what it can perceive at the outset may turn out to be rather limited.

The relation that makes the biggest difference in a (partially) lawful world, yet are compatible with cumulative learning, are the rules that associate an action to an effect. Anything that reliably can be done to affect the world in particular ways – with primitive actions, by the way, because there are limits to how much sophistication should and can be stored in a seed – could serve as a starting point for learning about the world. Learning to achieve goals in a complex world can only be achieved by causing something to happen, but to cause something to happen – out of infinite possibilities – means we must know about causal relations. Causal relations capture in a compact way a principle of how things come to be: Via interactions between demarcated phenomena, where one begins and triggers another. This is the most fundamental assumption that any general learner can make: *Assumption of causal relations—ACR:*

§2 *Causal relations must be a fundamental organizing principle of general knowledge representation.*

Others have argued the same (Pearl 2012) from a different angle, but for similar reasons. Note that we do not have to answer any philosophical questions about whether causes and effects "really exist" or whether they only exist from the viewpoint of the observer—this is simply what we call the fact that in a lawful world, even a *partially* lawful one, some things are better predictors of particular outcomes than others. It is a practical definition, and very fitting as a fundamental principle for a practical solution to practical problems (i.e. intelligence).

Also, if the agent is to learn some aspects of perception, we might want to put some general (meta) rules about extracting patterns from the world (this is in effect what nature does in many species; in effect this is a mapping from expected world stimuli to the agent's percepts). These, too, should be made as general as possible for a general learner, but on the downside it limits the range of other rules we may put in the seed because to learn, the agent must be able to perceive from the outset.

Another conclusion from this is that the seed must make *some assumptions* about the world; it cannot offer a completely blank slate – a tabula rasa – because without reference to *something observable* in the world, the newborn learner has nothing to grab onto.

§3 *The seed of a newborn autonomous general learner must reference variables that it can affect and sense.*

From these beginnings, models of causal (and other) relations would be built (Figure 2). However, this grounding cannot proceed unless there is a causal relationship between the seed variables (actions and observables).

If this were all there is to the story we would simply stop here. To work with causal relations through cumulative learning requires mechanisms to deal with the passage of time, memory fill-up, a potentially exponentially growing set of rules generated from a broad range of experiences, and an inevitably increasing set of contradicting information.

5. Cumulative Learning

Any controller that consistently, effectively, and efficiently achieves its goals is a 'good controller'.[14] As shown by Conant and Ashby's *Good Regulator* theorem (1970), every good controller of a system must contain a model of that system. Models, in our approach, are autonomously produced via a process that relies on contrasting prior knowledge (or seed) and percepts of observed variables and their relations in the world, supplied from experience. This process is really an semi-experimental empirical process, equivalent to the scientific method (which is also model-based (cf. Dellsén 2018)). Any such process, in an intricate task-environment, faulty models are inevitable. To evaluate and verify models, they are measured on their *usefulness* for achieving goals: The more accurately, effectively and efficiently they help the controller achieve its goals and sub-goals, the more useful they are. A cumulative learner is thus also a cumulative modeler, and its capacity to learn lies in its combined ability to create models and use them.

Following Thórisson et al. (2019), we consider a cumulative learner a *learning controller* that is constantly encountering novelty that it must *unify* with its current knowledge—it is an *experience-based* learner (Steunebrink et al. 2016, Wang 2006a). The learner's behavior, including its learning, is guided by one or more top-level drives (a drive hierarchy), that results in regularities being extracted recursively from its experience – of self and environment – to construct unified knowledge useful for achieving goals in that environment (Thórisson and Talbot 2018b).

Goals of varying complexity may be assigned to a goal-driven learner, by itself or someone else, composed of a set of sub-goals, each possibly involving a relatively large set of variables spanning potentially long periods of time, whose successful achievement requires diligent tracking of time, at multiple orders of magnitude. Goals, relations, variables, and other aspects of the world are represented as a growing network of (micro-) models.[15] Cumulative learning is a process of *model unification:*

14. Points of reference for these adverbs can more or less come from any source; in nature some obviously come from starvation, death, and procreation.
15. In prior work we have referred to such models as 'peewee-granular' models (Thórisson 2012, Thórisson and Nivel 2009).

> *New information enters by default into a process of being integrated with already-acquired knowledge—whether it is in agreement with it or not. This is compression under requirements of* incrementality, realtime, *and* generalization: *Replacing incorrect knowledge and extending current knowledge frequently, while generalizing when possible, prepares knowledge to be efficiently applicable to the largest possible class of situations, tasks, topics, and domains—as soon as possible during the learner's lifetime.* (Thórisson et al. 2019, p. 199)

5.1 Cumulative Modeling

When brought into an unknown world with only a bootstrapping seed referencing a handful of variables that may or may not be causally related in those particular circumstances, one must stay alert to indicators of causal relations. In the "worst case scenario"—when a set of percepts seem as close to completely novel as they get—the only resort is working from *correlations*: The co-occurrence of events. Of course, correlation does not mean causation, but no causation is without correlation (as long as some relevant variables are observable), which is why this is the most general way to bootstrap world modelling (*CRM*s) from meager beginnings.

§4 *An autonomous general learner bootstraps its knowledge about novelty from observed correlation.*

A *CRM*s may be created when the controller observes an event α and a subsequent event β that follows. The model can be seen as a hypothesis that the observed event α caused the observed event β, so that when observing again en event α in the future, this model will predict that β will be observed. Models that do this prediction better than others are kept and used, others are deleted. When a better model comes along it will be preferred over the old one(s).

Large dynamic state spaces (such as the physical world) will of course present, for any period of time of interest to a learner, a large amount of semi-co-occurring events. A practical rule of thumb, to limit the search space for interested relationships, is to shrink down the spatio-temporal scope of such a search (it is no coincidence that Pavlov's dogs salivate more when the bell during training is placed closer to feeding time, following specific spatio-temporal curves (Pavlov 1902; a similar thing is found in summing of inputs in natural neural networks).

A cumulative learner's hope of successfully unifying new models with its current knowledge is based on using current knowledge as a kind of seed, whereby new models is contrasted with current knowledge so that it may help highlight similarities and differences, through a comparison process. A newborn with little knowledge (just its seed), and limited acting and sensing abilities, may be facing the most extremely constrained knowledge acquisition in its lifetime. Yet the two situations are identical in that a seed that is effective, as opposed to a less effective one, will help a learner with *new information*. That is in fact what *any* useful knowledge will do – according to the canonical definition of learning – and therefore we consider

§5 *experience-based seed-programmed bootstrapping a special case of general cumulative learning.*

For all intents and purposes, this approach to modelling applies equally to seed-based programming as to model-driven cumulative learning in general. If the new information is only a small missing detail in an existing fairly complete model set (i.e. the agent already knows a lot about the phenomenon that the detail belongs to), unification may be straight forward because there are no conflicts with existing models and the new information simply "snaps in place." For any agent facing a situation or phenomenon whose knowledge about that is already advanced, this is how learning will proceed. The difference between two tasks, $f_\Delta(task_1, task_2)$, may be measured by an agent's ability to predict things in that domain.

For modeling the world around it, an autonomous learner in the physical world is not just limited to its own exclusive experience, it is also limited to *incremental* information collection. To make new information available as soon as possible for action and planning, and to avoid (spatio-temporal) *semantic siloing*, unification should happen as soon as possible. If the information "chunks" to be unified are too big, siloing can happen due to a combinatorial explosion, making unification in essence impractical, resulting in either partial or incorrect unification, or complete abandonment of the effort. Part of the difficulty lies in the inevitable verification of the unification in an uncertain environment, where computational power, time and energy are limited and axiomatic proof is impossible to come by.

A practical alternative is *bounded* recursive self-improvement (Nivel et al. 2013b) whereby unification is accomplished by frequent and small increments, relative to the learner's lifetime. This is as obvious in the simplest of real-world cases as it is in more involved ones, whether it is when we quicken our pace when we're about to miss the bus, or start to save money a year before our scheduled expensive world cruise. By extension, the models should also be semantically simple (have few parts with simple operational principles).[16]

As discussed by Steunebrink et al. (2016), due to semantic (spatio-temporal) siloing, self-modifications must be fine-grained, tentative, additive, reversible, and rated over time as experience accumulates—concurrently with all other activities of the system. The relatively simpler semantics of comparing small "snippets" of experience to existing knowledge than large ones means unification becomes more manageable. This is essential in worlds with limited time and energy, LTE. The comparison proceeds by isolating similarities and differences ("micro-analogies"), that allow new percepts to be dissected into parts and their relations. Because such dissection involves both static and dynamic properties, the process is a multi-dimensional comparison of dynamic models. An intermediate result of this process is a set of relevant known related knowledge, on which hypotheses can be based, that direct exploration of their implications using deduction-based simulations over established knowledge (good models). When simulated implications contradict current knowledge, choices must be made for how to resolve them.

One challenge a cognitive controller faces is finding the best balance between identifying and retaining (in memory) only variables that matter to the task at hand and other related variables that (e.g. in the future) might interfere or help with tasks/goals—*relevant* information. Spending a lot of time on identifying and modeling apparently relevant variables may end up being wasted time, should those models never be needed. Nevertheless, even in an environment with a large range of variables spanning broad variability, sufficient expe-

16. This has the added benefit of making model behavior during processing highly predictable with respect to time, memory, and other resources, which is critical for recursive planning (where planning time is taken into account when planning (Nivel and Thórisson 2013b, Nivel and Thórisson 2008)).

rience with the relevant variables on separate instances of tasks undertaken should enable a good modeler to make models that allow it to predict and achieve goals with increasing accuracy and proficiency. As the quality of models increases, the frequency of surprises are reduced, and controller performance compromised by probability factors is lowered.

In our conceptualization, similarity and relevance are close cousins: Those models that share variables with a current situation and relevant goals are highly likely to be relevant, because they talk about how some patterns in the current situation are likely to be transformed over time, as well as say something about how the controller can achieve transformations that result in goal achievement. We will come back to this in Section 5.4.

5.2 Semantic Modularity & Operational Closure

From the initial creation of new knowledge, and throughout its subsequent modification, expansion, unification, deletion, and usage, an agent's learning mechanisms must be capable of self-supervised "surgical" operation on existing knowledge, as new evidence comes to light through experience. The information that makes up such an agent's knowledge set must be structured in a way that supports reflective processes including discrimination, comparison, and manipulation of *arbitrary subsets of the knowledge set*, so that e.g. over-generalization may be tightened, regularities shared by disjoint phenomena unified, recurring patterns generalized beyond their instances, new relations $\Re$ created, deleted and changed, counterevidence weighed, conflicts ironed out, counterfactuals and alternative outcomes considered, and so on.

This kind of incremental "surgical" knowledge construction requires that the system knows (a) *what* existing knowledge the new information is *relevant* for, (b) *how* relevant it is to that knowledge, and (c) *how to unify* it in the right way. As we have established, the models representing the new information must be small and transparent. This is done through comparison, for which we may define a similarity comparison operator (Sheikhlar et al. 2020). Additionally, the new information must be represented in a format compatible with the existing knowledge, lest such comparison becomes impractically complex of computationally expensive.

A knowledge representation that meets these requirements is *semantically modular*, meaning that subsets of the knowledge carry white-box spatio-temporal semantics, such that that the application, relevance, and structure of the knowledge is retained when dissected. This semantic modularity must carry down to the smallest knowledge "nugget" that the system intends to operate on when unifying new knowledge with old. Otherwise the knowledge becomes *siloed*, building up over time as separate, loosely or entirely unconnected "buckets" of knowledge, each bucket being relevant to a particular and subset of the task-environments the learner has encountered. When opportunity arises to update the knowledge, each bucket – being independent – can only be updated in its entirety. In this process the challenge is to ensure the maintenance of the validity of prior knowledge. This is increasingly difficult to do the larger the silos are. The worst-case scenario is a single silo with no semantic modularity, which must be updated in its entirety when new information appears relevant, no matter how small. Silos should be avoided because, in addition to being inefficient, they presents a heightened risk of regression errors, resource shortage, and incompatibility with original top-level goals. Ashby's *Principle of Requisite Variety* (1958), "AI's corollary to Nyquist's Sampling Theorem," states that the resolution of a system's representation must be at least as fine-grain as the smallest detail that it intends to handle. It lies at the core of our argument for semantic modularity and transparency.

Whether a CRM set refers to an atomic (relatively) well-defined detail such as the color of a jacket or a compound ill-defined concept like "working in Iceland," it must contain information that uniquely demarcates its target reference. This can be done by a particular and unique pattern of relations to other models, with which it shares some patterns, which in turn requires a similarity comparator (like the Ψ operator in Sheikhlar et al. 2020). Rather than forming a rigid structure, like classical ontologies (cf. Chandrasekaran et al. 1999), such patterns may be computed on an event-driven just-in-time basis, based on the current situation (as grounded through percepts—world situatedness) and active goals (as grounded in drives—cognitive situatedness). Any atomic concept (many of which are closely tied to a basic function of transducers, e.g. color), may require only a few models to represent, in the simplest case; any complex concept will consist of a network of models, which again, are dynamically created based on the two "grounding sides" of the "cognitive *now:*" Situatedness and drives.[17]

As others have pointed out (cf. Wang 2006a), a system that learns from experience in a constructivist manner – i.e. constructs its own knowledge from experience – must build its mental constructs on a meaningful relationship between thought and effect, and this, we have already laid down some arguments for, boils down to modeling cause and effect. By "effect" we mean changes in variables measurable by a controller's transducers, but also traceable (internally) by the controller's own monitoring of its own operation. The meaning of any mental operations – its *operational semantics* – is thus defined by its relation to monitored, operationally accessible and semantically discernible events. To be of any use to learn (to get better at targeted tasks, as well as getting better at getting better), these operational semantics must be *penetrable* by the system itself—their mechanisms must fit with mechanisms available to the controller. The cyberneticians called this "closure" (Ashby 1956). In our approach, for mental operations to achieve semantic and operational closure with respect to the physical world, auto-generated CRMs evaluated in-situ via transducers, coupled with forward and backward chaining, are used. For mental development to achieve semantic and operational closure, reflection is needed.[18] Both are managed through event-driven auto-catalytic processes. Because the cognitive processes are sufficient for giving the system the ability to create, evaluate, manage and update knowledge created this way, it has informational closure. These principles informed the implementation of our Autocatalytic Endogenous Reflective Architecture (Nivel et al. 2013c).

17. Because our CRMs are fine-grain, any complex invariant in the physical world that needs to be computed – e.g. the apparent color permanence of a berry in in direct sunlight and in the shadow (which will register differently in the transducers but which the controller will want to perceive as being the same) – may require a large set of CRMs. Not all relevant models for such computation, however, need to be used on every occasion—this will be determined dynamically by context.

18. The distinction is not necessarily as perfectly sharp as this may sound: Either may be useful for the other purpose.

5.3 Causal-Relational Models

An atomic model, in our approach, is based on relationships between two patterns, one describing prior condition and the other a subsequent condition (Nivel et al. 2013b). The prior state we say is on the left-hand-side (LHS) of the model, the subsequent on the right (RHS). Such models capture transformations of situations (states, events, conditions) which are represented as patterns referenced by variables, states, or other models. Read from left to right, a model is thus a predictor that says "when I see the LHS pattern I hypothesize RHS will come next." The predictions come with a time attached.

Models M_Φ is a set containing models of a phenomenon Φ $\{m_1 \ldots m_{\|\mathcal{M}_\Phi\|} \in \mathcal{M}_\Phi\}$. The closer the information structures $m_i \in \mathcal{M}_\Phi$ represent elements (sub-parts)[19] $\varphi \in \Phi$, at any level of detail, including their couplings $\Re_\Phi$, the better a system with models $\mathcal{M}_\Phi$ can predict Φ. For a model created from scratch based on correlation between two patterns, α and β, a strict requirement is that any pattern assumed for the LHS must meet a condition of having be measured to exist *prior to* the RHS, $\alpha(t_1) < \beta(t_2)$. This is one way to have the bootstrapping of learning start out honing in on causal relations from the very beginning. A model m's *reliability* can be assessed at any time t:

$$(m, t) = \frac{e^+(m,t)}{e(m,t)+1}$$

where $e^+(m,t)$ is the number of successful predictions produced by m and $e(m,t)$ the total number of predictions, both updated each time a prediction fails or succeeds (Nivel et al. 2015).

In addition to be used for prediction, models can be used for planning, because they are bi-directional; the very same model that predicted "if α then β" when read left to right, should say when read the other way, "if I want the RHS pattern β, I should look for the LHS pattern α." This means that LHS patterns function as sub-goals, making every LHS in a chain of models a potential step in a plan.

The *context* of a phenomenon Φ is defined by its outward relations, $\Re_\Phi^{out}$. For any complex phenomenon in a complex world, completeness of $\Re_\Phi^{out}$ is generally not to be expected, as this may be an extraordinarily large number. However, for any two phenomena Φ_1 and Φ_2 that are related, if $\|\Re_{\Phi_1}^{out} \cap \Re_{\Phi_2}^{out}\| = small$, then predicting $\Re_{\Phi_1}^{out}$ may not require a lot of models of Φ_2, even if $\|\Re_{\Phi_2}^{in}\| = large$. An agent whose models are only accurate for $\Re_\Phi^{in}$ can predict Φ's behavior in "isolation" but not how it interacts with other phenomena; if models are only accurate for $\Re_\Phi^{out}$ the agent can predict Φ's relation to other phenomena, and thus how it interacts with them, but will not be able to predict the behavior of Φ's internals, forcing a black-box view.

How these *CRM*s are created from observation is not obvious; we will now describe how our combined abduction-deduction methodology removes less causal models based on correlation out of the model set $\mathcal{M}$, increasing the ratio of causally useful models over time.

19. By "elements" and "sub-parts" we mean any sub-division of Φ, including sub-structures, component processes, whole-part relations, causal relations, etc.

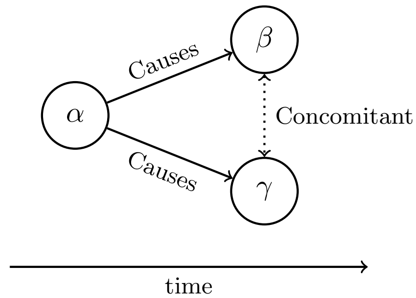

Figure 3: Relations between three variables α, β and γ
(from Thórisson and Talbot 2018b).

5.4 Micro-Ampliative Reasoning

As we have seen, predictions through deduction are a central method for achieving self-supervised learning in our theory: Using available CRMs to anticipate what lies ahead. While prediction with CRMs is similar to the logical implication ($\Rightarrow$) in that in the *perfect* case a model that states $\alpha \Rightarrow \beta$, i.e. that if you see α you will see β, will always be true. However, our models model non-axiomatic data, based on the experience of an agent over time (e.g. in the physical world). Our deduction therefore cannot be *classical axiomatic* deduction, because the physical world's axioms will be forever out of reach; the best a learner in this situation can do is hope that what it has come up with so far in terms of generalizations, is *good enough*. This means that even if we had a CRM m_1 that says $\beta \rightarrow \alpha$, and experience tells us this always to be the case, there is still no guarantee that α really is the cause, because that is only based on our experience *so far*. In terms of real causal relations, these models are therefore rightfully considered *pragmatic hypotheses* of causal relations, standing as potential indicators for what *might be the case*—and that's why they need to be revisable, augmentable, verifiable, and contextualizable. In addition to deduction, this may be done through abduction, induction and analogy—ampliative reasoning.[20] We will first look at abduction, then analogy (induction will not be addressed in detail here).

Consider a system where a cause α has two effects, β and γ (figure 3). We assume that to the agent, α appears before β and γ, but β and γ appear together. Four models can be used to describe what the agent experiences every time it observes these variables:

$$m_1[\beta \rightarrow \gamma] \quad m_2[\gamma \rightarrow \beta] \quad m_3[\alpha \rightarrow \beta] \quad m_4[\alpha \rightarrow \gamma]$$

Any of these models will predict observed events correctly: If you observe β you will observe γ, and vice versa; if you observe α you will observe β and γ. They can be combined to cover the full experience with all variables: m_1 and m_3; m_2 and m_4; m_3 and m_4. However, not all of them represent the actual causal relations, and not all of them can be used to achieve goals in the α-β-γ domain. Let's say that α is a light switch, β is a light bulb, and

20. Peirce's use of the concept of 'ampliative reasoning' included abduction, induction and analogy (cf. Psillos 2011); ours adds (corrigible) deduction to that list.

γ is the room appearing when lit up by the light. If When you press α, the light turns on. When you press α again, it turns off. If you want to stop γ from appearing it does not help to manipulate β, or vice versa (you could break the lightbulb, that is beyond the scope of this example)—to manipulate either β or γ the only variable that will help achieve the right sub-goal is α, as specified by m_3 and m_4.

Models that do predict correctly, however, may still be wrong about the causes of the predictions; i.e. the LHS of a model might not be an exclusive cause of the RHS (for the LHS to be an effective[21] necessary and sufficient cause of the RHS in conditions Z, the RHS must disappear in the absence of LHS and appear in the presence of the LHS, given no changes in Z). A good predictor that does not appeal to (semi-) deterministic necessary-and-sufficient causes is no good for getting things done as it cannot be relied on for effective action (we remind the reader of the New York Mayor who banned summer ice cream sales in Central Park to stop muggings, because the evidence showed strong correlation between the two). To create models that are not only good for prediction but also for getting things done, backward chaining – reading a CRM from right to left – tells us how to use a model to make a plan: If the RHS is a goal, or part of a causal chain constituting a potential plan, the model's LHS is a sub-goal *if and only if it corresponds to an effective cause in the world it models*. This is in effect abduction.

When each of these models, m_1, m_2, m_3 and m_4 is used for *both* prediction and goal achievement, models m_1 and m_2 will be deleted due to their incorrect predictions: It does not help to manipulate either β or γ directly to get γ to appear. What remains is a model $m_X : [fact(light = off) \wedge fact(press(\alpha)) \rightarrow fact(light = on)]$ that tells the agent that "buttons can turn on lights".[22] Repeating such operations will leave only models that capture useful causal relations in the domain, to the extent that this can be represented as relationships between *observable* variables: If the agent wants the room to appear it presses the button using model m_X.

Once candidate models – really, hypotheses of how the world hangs together – have been created, they might predict incorrectly, in which case they will get an increasingly lower score over time and soon be deleted (they will also be deleted if other models exist that work better). Through the creation of a number of such models, with patterns on each side referencing internal as well as external variables, and testing the models through a process of deduction and abduction, a network of *useful* models is built up. Notice that a model that can be relied on to get things done will *also* be good for prediction—i.e. we do not lose the predictive power of the model set through this pruning, we simply remove those that are no good for planning and guiding action. Such processes are explained further in Thórisson and Talbot (2018b). We conclude that:

§6 *Unified abduction and deduction centered on a single model can isolate causal relations in experience data.*

21. In the limit an "effective" identified cause is deterministic; in the physical world it is sufficient to find factors that are useful because they work most of the time.
22. More specifically, the generality of the model, e.g. whether it reads "this button will turn on that light" or "buttons of this kind tend to turn on lights of that kind" will be determined by how the LHS and RHS patterns are represented. There are other aspects which we omit for clarity, including information about time, context, and transformation functions of variable values for both deduction and abduction. Details can be found in Nivel et al. 2015 and Nivel et al. 2013c.

For a newborn, as well as a learner with prior partial knowledge of a phenomenon or situation, the actions it takes must have a realistic potential to result in changes that are observable to it. What must be learned in this knowledge bootstrapping is information that supports both prediction and planning, because otherwise the learner cannot use learned predictions to affect the world—which, after all, is a fundamental reason for intelligence. In other words, pragmatic knowledge is fundamental to knowledge bootstrapping. Thus, a necessary job of an autonomous general learner is to reverse-engineer causal chains. We further hypothesize that this process lies at the heart of all novelty learning—that when faced with truly novel phenomena or in truly novel situations, an autonomous general learner will rely on these principles. We summarize this as:

§7 *Extraction of causal relations from experience data is fundamental to knowledge bootstrapping.*

With a lot of models, some – possibly most – will be irrelevant to any particular task or subtask at a particular time. For this there must be a way to find the most relevant ones. We define a computing similarity between patterns, Ψ, a multi-dimensional comparison computation using ampliative reasoning (Sheikhlar et al. 2020), that takes two or more patterns and returns their similarity.[23] Running $\Psi(\vartheta, \mathcal{M}, \mathcal{G})$, where ϑ is a target pattern (from a model, a goal, percept, or other source), $\mathcal{M}$ is a set of models and $\mathcal{G}$ a set of goals (goals are also defined by patterns), produces a similarity gradient over them that serves as an indicator of relevance, because only if the models share (subsets of, or entire) patterns may they reference shared phenomena. If we define the *aspects* of a phenomenon as a subdivision of it, $\mathcal{A} \subset \Phi$, that is of pragmatic importance to the agent's goals and tasks; every aspect is verified through multiple percepts predicted by relevant models. Aspects that an agent can predict this way would be familiar to the agent, $\mathcal{A}_{fam}$; those it can't predict are novel, $\mathcal{A}_{nov}$ (Sheikhlar et al. 2020).

In Sheikhlar et al. (2020) we describe a multi-dimensional method for computing analogies, employing the above principles. With the addition of ways to narrow the *relevant* models, patterns, aspects, and percepts, the similarity computations can be used to create analogies at any level of detail. On the basis of further similarity computations, it can proceed to (a) dissect the novel aspects and generate new models, possibly patterned after similar yet incorrect models, in an attempt to (b) fill in missing knowledge gaps. To summarize this principle:

§8 *Bootstrapping on variables that are potentially causally related is necessary for all novelty learning.*

As already discussed, this is done by observing correlation and creating CRMs based on them.

6. Summary: Seed-Programmed Autonomous General Learning

In a world where rule A sometimes holds for situations of type T and rule B for other times, and A and B are in some ways incompatible, this may mean that something is amiss—either

23. This comparison function is implemented trough several mechanisms in our cognitive architecture; its description is simplified here for clarity.

the models of these rules are incorrect, the observations on which they are based, or the models are missing some facts. A search for a variable that allows predicting when either A or B should be used may lead to more effective models. A world with hierarchical rules allows systematic search for these using logic, even if the rules must be inferred from observations of changes in variables over time.

As informationally rich open-ended complex environments cannot be axiomatized beforehand; very few assumptions can be made about the knowledge that a system may have to acquire in such an environment: The information available to be perceived and the behavioral challenges that the system's future environment(s) may present, may be of many types, and the specifics of these types cannot be known beforehand. The generality of the system is thus constrained by its ability to deal with this potentially large set of information, as well as its own skills, in an organized manner: the less specific to particular types of information the knowledge representation is, the greater the system's potential for being general. All observations are by definition specific; an efficient way to generalize them is through reasoning (not just induction—we have seen that induction must be accompanied by other forms of reasoning to be of any use).

To rely on logic means doing reasoning. In our theory reasoning permeates the knowledge creation and management processes from a very low level of detail and upwards, using all available forms of (ampliative[24]) reasoning—deduction, abduction, induction and analogies. To emphasize the fine-grain level of detail, and the contrast with higher level of (human-like) conscious reasoning, we call them *micro-ampliative* reasoning methods. Micro-ampliative reasoning supervenes on the fine-grain causal-relational models described above, ensuring consistency at a very fine-grain level, based on similarity (proximity in analogy space).

In short, the ampliative reasoning methods, and their purpose, are as follows (the list is nothing new but the way we employ them is):

- ☐ Deduction, to produce plausible turn of events: *Simulation of causal chains from a given state.*
- ☐ Abduction, to find plans that work: *Constraint-based backward-chaining from a given end goal through plausible causal chains that can lead to it from a given starting state.*
- ☐ Induction, to generalize observations: *Assuming general rules first, then reducing scope only when evidence requires.*
- ☐ Analogies, for discretionary comparisons of knowledge and percepts: *Assuming new rules have maximal generality, then reducing scope only when evidence requires.*

The kind of reasoning we are talking about is a kind of "micro-reasoning": Fine-grain ("peewee") causal-relational models are continuously organized via these processes, and therefore no major (dangerous or practically prohibitive) overhaul of the knowledge at the smallest-grain representation is needed Thórisson and Nivel 2009).

To improve incomplete and incorrect knowledge with increased experience – accumulated evidence – knowledge acquisition processes bring already-acquired knowledge opportunistically (but systematically) and constantly towards greater coherence and consistency. For

24. Peirce's suggested 'ampliative reasoning' to include abduction, induction and analogy making Psillos 2011; we include (corrigible) deduction as well.

the initial creation of knowledge, and its subsequent usage, expansion, modification, unification, and deletion, construction mechanisms are self-guided, capable of self-supervised "surgical" operation on existing knowledge, through self-inspection or *reflection*.

To allow appropriate application to any situation the knowledge must be compared to observed variables in the environment, which means it must also be possible to selectively compare two or more knowledge structures. Flexible and selective comparison means in effect that the knowledge supports analogies. If the learner is to get better at making analogies this means the analogy making mechanisms must also be comparable: "For situation X and purposes Y, analogies of type A work better than of type B". This means that the learner must be capable of *introspection*.

An implementation of this general scheme was used in our Autocatalytic Endogenous Reflective Architecture (AERA; Nivel et al. 2013c, Nivel et al. 2013b, Nivel and Thórisson 2013b), in a way that allows self-modeling: Since the models can reference each other, via their LHS and RHS patterns, any plan the system creates can involve the internal planning mechanisms (e.g. which models are used during the plan) as well as how to act with respect to the environment—while the system keeps track of the referents of each of the models, the mechanism itself does not care what the references are, in essence creating a "syntax for thinking." During a normal run of AERA, a multitude of models are being used simultaneously to continuously predict and plan, in parallel.

In one evaluation setup of AERA, the S0 agent was given the task to learn how to meet the demands of a human telling it to move one of four objects around on a (virtual) tabletop. AERA received a stream of data generated from the computer graphics and speech. The speech included a few phrases (e.g. "Put that there," "Now put it here," "Take the blue cube ... and put it next to the red sphere"). There were four objects, two cubes, blue and red, and two spheres, also blue and red. The seed of S0 contained one goal (to please the human, operationalized as seeing the words "thank you," a small ontology telling it the names of the objects, and 6 CRMs. It had no knowledge of grammar or syntax, deictic (pointing) gestures or what words referring to placement meant. After observing two humans playing a game of "put that there" for about 2,5 S0 had created 20 models of its own, for a total of 26 CRMs that allowed it to perform the task flawlessly in interaction with a human.

In another evaluation scenario, our AERA-S1 agent was given a seed to learn to conduct a TV-style interview. The interview was conducted between the agent and a human participant in cyberspace—akin to a video conference. S1 learned by watching two humans in an interview between an interviewer and an interviewee who was an expert in recycling. No information was given to S1 about syntax, turntaking, deictic gestures, anaphora, or how to construct an answers to questions; the seed contained only five drives, 26 models and a small ontology of graphical elements. After 20 hours observation, S1 could perform the interview with a human, in either role of interviewer or interviewee, flawlessly (Nivel et al. 2014c).[25] After the 20 hours it had auto-created 1400 CRMs that constituted its own rules for syntax that completely covered the language use it had observed, including some generalizations, so that it knew how to direct the interviewee's attention to objects using three different deictic methods (picking up objects, pointing to them with an index finger

25. Descriptions of the architecture can be found in Nivel et al. (2015), Nivel et al. (2014b), Nivel et al. (2013b), Nivel and Thórisson (2013b) and Thórisson and Talbot (2018a); results of experiments with S1 are detailed in Nivel et al. (2014c) and Nivel et al. (2014d).

and waving a hand in their direction), as well as use such information to infer which objects were meant by "it" and "this." It could also handle continuity, as when the interviewer said "tell me more."

7. Conclusions

Our aim is to create general machine intelligence (GMI). To get there requires meeting a number of requirements, some of which are generally agreed upon in the research community (cf. Thórisson et al. 2015, Laird and Wray III 2010), others which undoubtedly are less obvious and therefore more controversial. Due to the requirement of such GMI having to eventually address the physical world, we have argued for a non-axiomatic approach to autonomous cumulative learning. We have presented ideas for how this may be achieved. A major thrust of the arguments in favor of our approach comes from the need for semantic closure and semantic compositionality: Representations that carry operational semantics down to fine-grain "peewee" levels, allowing "surgical precision" modification and updating that implements autonomous cumulative learning. This follows from the fact that learning in the physical world is always incremental ("life-long") and large updates of the knowledge, when new (and sometimes contradicting information) comes to light, will be too expensive to be practically implementable. Our proposed approach enables autonomous bootstrapping from existing knowledge, allowing bootstrapping of novel information. The results from implementing this general approach in two systems, AERA (Nivel et al. 2013c) and NARS (Hammer and Lofthouse 2020, Wang 2006b), both of which have been demonstrated to handle unspecified task-environments and unknown transform functions only accessible over time via multi-dimensional observable variables, presents evidence that this approach is both viable and scalable. Their ability to use contexts not specified beforehand to help with their learning of novel spatio-temporal patterns further lends credibility to the approach. Future work involves further development of these principles, especially w.r.t. concept creation and higher-level communication.

Acknowledgments

I am deeply indebted to my friend and collaborator Eric Nivel for countless and extended in-depth discussions about these topics and related ideas over the past fourteen years. I would like to thank Bas Steunebrink, Jeff Thompson and Pei Wang, as well as my students Jordi Bieger, Arash Sheikhlar, Leonard Eberding, Helgi P. Helgason and Throstur Thorarensen for discussions about many of these subjects. Furthermore, I would like to thank members of the HUMANOBS team for their valuable insights (you know who you are!). Leonard Eberding provided useful comments on drafts of this paper. This work was supported in part by the Icelandic Institute for Intelligent Machines, Reykjavik University, and Cisco Systems Inc. —*KRTh*

References

Aristotle. *De Anima*. 350 BCE.

W. Ross Ashby. *An introduction to cybernetics*. Chapman Hall, 1956.

W. Ross Ashby. Requisite variety and its implications for the control of complex systems. *Cybernetica*, 1:83–99, 1958.

Dmitry Balakhonov and Jonas Rose. Crows rival monkeys in cognitive capacity. *Scientific Reports*, 7:2045–2322, 2017. doi: 10.1038/s41598-017-09400-0.

T. F. Beernaert, A. E. Bayrak, L. F.P. Etman, and P. Y. Papalambros. Framing the concept of autonomy in system design. *Proceedings of International Design Conference, DESIGN*, 6:2821–2832, 2018. ISSN 18479073. doi: 10.21278/idc.2018.0281.

Tarek R. Besold and Ute Schmid. Why generality is key to human-level artificial intelligence. *Advances in Cognitive Systems*, 4:13–24, 2016.

Jordi Bieger and Kristinn R. Thórisson. Evaluating understanding. In *IJCAI Workshop on Evaluating General-Purpose AI, Melbourne, Australia*, 2017.

Jordi Bieger, Kristinn R. Thórisson, and B. Steunebrink. Evaluating understanding. In *IJCAI Workshop on Evaluating General-Purpose AI, Melbourne, Australia (2017 to be accepted)*, 2017.

Jordi E. Bieger and Kristinn R. Thórisson. Task analysis for teaching cumulative learners. In *Proc. Artificial General Intelligence*, pages 21–31. Springer International Publishing, 2018.

Balakrishnan Chandrasekaran, John R. Josephson, and V. Richard Benjamins. Ontology of tasks and methods. 1999.

Roger C. Conant and W. Ross Ashby. Every good regulator of a system must be a model of that system†. *International journal of systems science*, 1(2):89–97, 1970.

Finnur Dellsén. Scientific progress: Four accounts. *Philosophy Compass*, 13, 2018.

Werner Dubitzky, Olaf Wolkenhauer, Kwang-Hyun Cho, and Hiroki Yokota, editors. *Encyclopedia of Systems Biology*. Springer Publishers, 2013.

Leonard M Eberding, Arash Sheikhlar, Kristinn R Thórisson, and Sindri P Andrason. Sage: Task-environment platform for evaluating a broad range of AI learners. In *International Conference on Artificial General Intelligence*. Springer, 2020.

David Garlan, Robert Allen, and John Ockerbloom. Architectural mismatch or why it's hard to build systems out of existing parts. In *Software Engineering, 1995. ICSE 1995. 17th International Conference on*, pages 179–179, 1995.

James J. Gibson. *The senses considered as perceptual systems*. SHoughton Mifflin, 1966.

Ben Goertzel. Artificial general intellgience. In *Scholarpedia*, volume 10, page 31847. 2015.

Patrick Hammer and Tony Lofthouse. 'opennars for applications': Architecture and control. In *Proceedings of the 13th International Conference on Artificial General Intelligence*, 2020.

Francis Heylighen and Cliff Joslyn. Cybernetics and second order cybernetics. *Encyclopedia of physical science & technology*, 4:155–170, 2001.

John E. Laird and Robert E. Wray III. Cognitive Architecture Requirements for Achieving AGI. pages 1–6, 2010. doi: 10.2991/agi.2010.2.

John E. Laird and Robert E. Wray III. Cognitive architecture requirements for achieving AGI. In *Proc. of the Third Conference on Artificial General Intelligence*, pages 79–84, 2010.

W.F. Lawless, R. Mittu, D. Sofge, and S. Russell. *Autonomy and Artificial Intelligence: A Threat or Savior?* Springer, 2017.

E. E. Lee. Computing needs time. *Commun. ACM*, 52:70–79, 2009.

Juan Carlos Letelier, María Luz Cárdenas, and Athel Cornish-Bowden. From L'Homme Machine to metabolic closure: Steps towards understanding life. *Journal of Theoretical Biology*, 286(1):100–113, 2011. ISSN 00225193. doi: 10.1016/j.jtbi.2011.06.033.

Allen Newell. *Unified theories of cognition.* Harvard University Press, Cambridge, MA, 1994.

E. Nivel, K. R. Thórisson, and et al. Autonomous acquisition of natural language. *IADIS International Conference on Intelligent Systems & Agents*, pages 58–66, 2014a.

Eric Nivel and Kristinn R. Thórisson. Self-Programming: Operationalizing Autonomy. In *Proceedings of the 2nd Conf. on Artificial General Intelligence*, 2008.

Eric Nivel and Kristinn R Thórisson. Towards a programming paradigm for control systems with high levels of existential autonomy. In *International Conference on Artificial General Intelligence*, pages 78–87. Springer, 2013a.

Eric Nivel and Kristinn R. Thórisson. Replicode: A constructivist programming paradigm and language. Technical RUTR-SCS13001, Reykjavik University School of Computer Science, 2013b.

Eric Nivel, Kristinn R. Thórisson, H. Dindo, G. Pezzulo, M. Rodriguez, C. Corbato, B. Steunebrink, D. Ognibene, A. Chella, Jürgen Schmidhuber, Ricardo Sanz, and Helgi Páll Helgason. Autocatalytic Endogenous Reflective Architecture. Technical RUTR-SCS13002, Reykjavik University School of Computer Science, Reykjavik, Iceland, 2013a.

Eric Nivel, Kristinn R Thórisson, Bas Steunebrink, Harris Dindo, Giovanni Pezzulo, Manuel Rodriguez, Carlos Corbato-Hernandez, Dimitri Ognibene, Jörgen Schmidhüber, Ricardo Sanz, Helgi P. Helgason, and Antonio Chella. Bounded recursive self-improvement. *Tech report RUTR-SCS13006, Reykjavik University – School of Computer Science*, 2013b.

Eric Nivel, Kristinn R Thórisson, Bas Steunebrink, Harris Dindo, Giovanni Pezzulo, Manuel Rodriguez, Carlos Corbato-Hernandez, Dimitri Ognibene, Jörgen Schmidhüber, Ricardo Sanz, Helgi P. Helgason, and Antonio Chella. Autocatalytic endogenous reflective architecture. *Tech report RUTR-SCS13002, Reykjavik University – School of Computer Science*, 2013c.

Eric Nivel, Kristinn R Thórisson, Bas Steunebrink, Harris Dindo, Giovanni Pezzulo, Manuel Rodriguez, Carlos Corbato-Hernandez, Dimitri Ognibene, Jörgen Schmidhüber, Ricardo Sanz, Helgi P. Helgason, and Antonio Chella. Bounded seed-AGI. pages 121–130, 2014b.

Eric Nivel, Kristinn R Thórisson, Bas Steunebrink, Harris Dindo, Giovanni Pezzulo, Manuel Rodriguez, Carlos Corbato-Hernandez, Dimitri Ognibene, Jörgen Schmidhüber, Ricardo Sanz, Helgi P. Helgason, Antonio Chella, and Gudberg K. Jonsson. Autommous acquisition of natural language. In *Proceedings of the IADIS International Conference on Intelligent Systems & Agents*, pages 58–66, 2014c.

Eric Nivel, Kristinn R. Thórisson, Bas R. Steunebrink, Haris Dindo, Giovanni Pezzulo, Manuel Rodríguez, Carlos Hernández, Dimitri Ognibene, Jürgen Schmidhuber, Ricardo Sanz, and others. Bounded Seed-AGI. In *Artificial General Intelligence*, pages 85–96. Springer, 2014d.

Eric Nivel, Kristinn R. Thórisson, Bas Steunebrink, and Jürgen Schmidhuber. Anytime Bounded Rationality. In Jordi Bieger, Ben Goertzel, and Alexey Potapov, editors, *Proceedings of The Eighth Conference on Artificial General Intelligence*, volume 9205 of *Lecture Notes in Artificial Intelligence*, pages 121–130. Springer, Berlin, Germany, 2015.

F.G.P. Patterson and W. Gordon. *Twenty-Seven Years of Project Koko and Michael*, pages 165–176. Springer US, Boston, MA, 2001. ISBN 978-0-306-47461-3. doi: 10.1007/0-306-47461-1_15. URL https://doi.org/10.1007/0-306-47461-1_15.

I. P. Pavlov. *The work of the digestive glands.* Griffin, 1902.

J. Pearl. *do*-calculus revisited. In Nando de Freitas and Kevin Murphy, editors, *Proceedings of the Twenty-Eighth Conference on Uncertainty in Artificial Intelligence*, pages 4–11, Corvallis, OR, 2012. AUAI Press.

Judea Pearl. Bayesianism and causality, or, why I am only a half-Bayesian. volume 12, pages 19–36, Corvallis, OR, 2001. Kluwer Academic Publishers.

Judea Pearl. *Causality: Models, Reasoning and Inference.* Cambridge University Press, New York, NY, USA, 2nd edition, 2009. ISBN 052189560X, 9780521895606.

Jean Piaget. *The origin of intelligence in the child.* Routledge & Paul, 1953.

John L. Pollock. Defeasible reasoning and degrees of justification. *Argument and Computation*, 1(1):7–22, 2010.

Stathis Psillos. An explorer upon untrodden ground: Peirce on abduction. In *Handbook of the History of Logic*, volume 10, pages 117–151. Elsevier, 2011.

H. Putnam. *Reason, Truth and History.* Cambridge University Press, 1981.

Darryn J. Reid, M. J. Pilling, and Vladimir Ivancevic. *Mathematics Of Autonomy: Mathematical Methods For Cyber-physical-cognitive Systems.* World Scientific, 2018.

Ken Richardson, editor. *Models of Cognitive Development.* Psychology Press, 1998.

Luis Mateus Rocha. Syntactic Autonomy: Why There is No Autonomy Without Symbols and How Self-Organizing Systems Might Evolve Them. *Annals of the New York Academy of Sciences*, 901(1):207–223, 2010. doi: 10.1111/j.1749-6632.2000.tb06280.x.

Arash Sheikhlar, Kristinn R Thórisson, and Leonard M Eberding. Autonomous cumulative transfer learning. In *International Conference on Artificial General Intelligence*. Springer, 2020.

Linda Smith and Michael Gasser. The development of embodied cognition: Six lessons from babies. *Artificial life*, 11(1-2):13–29, 2005.

Bas R. Steunebrink, Kristinn R. Thórisson, and Jürgen Schmidhuber. Growing recursive self-improvers. In *Proceedings of Artificial General Intelligence*, pages 129–139, 2016.

A. Taski. The semantic conception of truth. *Philosophy and Phenomenological Researc*, 4: 341–375, 1944.

Kristinn R. Thórisson. Modeling multimodal communication as a complex system. In *Modeling Communication with Robots and Virtual Humans*, pages 143–168. Springer, 2008.

Kristinn R. Thórisson. A New Constructivist AI: From Manual Methods to Self-Constructive Systems. In Pei Wang and Ben Goertzel, editors, *Theoretical Foundations of Artificial General Intelligence*, volume 4 of *Atlantis Thinking Machines*, pages 145–171. Atlantis Press, Amsterdam, The Netherlands, 2012.

Kristinn R. Thórisson. Machines with Autonomy & General Intelligence: Which Methodology? In *Proceedings of the Workshop on Architectures for Generality and Autonomy 2017*, Melbourne, Australia, August 2017.

Kristinn R Thórisson. Discretionarily constrained adaptation under insufficient knowledge & resources. *Journal of Artificial General Intelligence*, 11:7–13, 2020. doi: 10.2478/jagi-2020-0003.

Kristinn R. Thórisson and Helgi P. Helgason. Cognitive architectures & autonomy: A comparative review. *Journal of Artificial General Intelligence*, 3:2:1–30, 2012.

Kristinn R. Thórisson and Eric Nivel. Achieving artificial general intelligence through peewee granularity. In *Proc. of the 2nd Conf. on Artificial General Intelligence*, pages 220–221, 2009.

Kristinn R. Thórisson and Arthur Talbot. Cumulative learning with causal-relational models. In *Artificial General Intelligence*, pages 227–237, Cham, 2018a. Springer International Publishing.

Kristinn R Thórisson and Arthur Talbot. Cumulative learning with causal-relational models. In *International Conference on Artificial General Intelligence*, pages 227–237. Springer, 2018b.

Kristinn R Thórisson, Jordi Bieger, Stephan Schiffel, and Deon Garrett. Towards Flexible Task Environments for Comprehensive Evaluation of Artificial Intelligent Systems & Automatic Learners. In *Proceedings of the International Conference on Artificial General Intelligence*, pages 187–196, 2015.

Kristinn R. Thórisson, Jordi Bieger, Throstur Thorarensen, Jona Sigridur Sigurdardottir, and Bas R. Steunebrink. Why artificial intelligence needs a task theory – and what it might look like. In *Proceedings of the 9th International Conference on Artificial General Intelligence*, pages 118–128, 2016a.

Kristinn R. Thórisson, David Kremelberg, Bas R. Steunebrink, and Eric Nivel. About understanding. In *Proceedings of the International Conference on Artificial General Intelligence*, pages 106–117, New York, NY, USA, 2016b. Springer-Verlag.

Kristinn R Thórisson, Jordi Bieger, Xiang Li, and Pei Wang. Cumulative learning. In *Proceedings of the 12th International Conference on Artificial General Intelligence*, pages 198–208, Shenzen, China, 2019. Springer.

Jane X. Wang, Zeb Kurth-Nelson, Dhruva Tirumala, Hubert Soyer, Joel Z. Leibo, Remi Munos, Charles Blundell, Dharshan Kumaran, and Matt Botvinick. Learning to reinforcement learn. *arXiv:1611.05763 [cs, stat]*, November 2016. arXiv: 1611.05763.

Pei Wang. *Rigid flexibility: The logic of intelligence.* Springer Science, vol. 34, 2006a.

Pei Wang. *Rigid Flexibility.* Springer, 2006b.

Pei Wang. From NARS to a thinking machine. *Advances in Artificial General Intelligence: Concepts, Architectures and Algorithms*, 157:75–93, 2007.

Pei Wang. Insufficient Knowledge and Resources-A Biological Constraint and Its Functional Implications. In *AAAI Fall Symposium: Biologically Inspired Cognitive Architectures.* Citeseer, 2009.

Pei Wang. *Non-Axiomatic Logic: A Model of Intelligent Reasoning.* World Scientific Publishing, Singapore, 2013. ISBN 978-981-4440-28-8.

Pei Wang. On the definition of artificial intelligence. 2019.

Juyang Weng. Developmental robotics: Theory and experiments. *International Journal of Humanoid Robotics*, 1(02):199–236, 2004.

Jason Winning and William Bechtel. Being Emergence vs. Pattern Emergence: Complexity, control and goal-directedness in biological systems. In S. C. Gibb, R. F. Hendry, and T. Lancaster, editors, *Routledge Handbook of Emergence*, pages 134–144. Routledge, 2019.

Michael Wooldridge and Nicholas R. Jennings. Agent theories, architectures, and languages: A survey. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 890:1–39, 1995. ISSN 16113349. doi: 10.1007/3-540-58855-8_1.