Vygotskian Autotelic Artificial Intelligence:

Language and Culture Internalization for Human-Like AI

Cédric Colas[1,*,+], **Tristan Karch**[1,2,+], **Clément Moulin-Frier**[1,3], **and Pierre-Yves Oudeyer**[1,3,4]

[1]INRIA Bordeaux – Sud-Ouest, Talence, France
[2]Université de Bordeaux, Talence, France
[3]ENSTA Paris, Palaiseau, France
[4]Microsoft Research, Montreal, Canada
[*]cedric.colas@inria.fr
[+]equal contribution

ABSTRACT

Building autonomous artificial agents able to grow open-ended repertoires of skills across their lives is one of the fundamental goals of AI. To that end, a promising developmental approach recommends the design of intrinsically motivated agents that learn new skills by generating and pursuing their own goals — *autotelic agents*. However, despite recent progress, existing algorithms still show serious limitations in terms of goal diversity, exploration, generalization or skill composition. This perspective calls for the immersion of autotelic agents into *rich socio-cultural worlds*, an immensely important attribute of our environment that is mostly omitted in modern AI, including deep reinforcement learning research. We focus on language especially, and how its structure and content may support the development of new cognitive functions in artificial agents, just like it does in humans. Indeed, most of our skills could not be learned in isolation. Formal education teaches us to reason systematically, books teach us history, and YouTube might teach us how to cook. Most importantly, our values, traditions, norms and most of our goals are cultural in essence. This knowledge, and some argue, some of our highest cognitive functions such as abstraction, compositional imagination or relational thinking, are formed through linguistic and cultural interactions with others. Inspired by the seminal work of the developmentalist Vygotsky, we suggest the design of *Vygotskian autotelic agents* able to interact with others and, more importantly, able to internalize these interactions within the agent so as to transform them into *cognitive tools* supporting the development of new cognitive functions. This perspective paper finds its inspiration in the work of psychologists and philosophers to propose a new AI paradigm in the quest for artificial lifelong skill discovery. It justifies the approach by uncovering several examples of new artificial cognitive functions emerging from interactions between language and embodiment in recent works at the intersection of deep reinforcement learning and natural language processing. Looking forward, it highlights future opportunities and challenges for Vygotskian Autotelic AI research.

Introduction

Humans are remarkable examples of lifelong open-ended learners. They learn to recognize objects and crawl as infants, learn to ask questions and interact with peers as toddlers, learn to master engineering, science, or arts as adults. A fundamental goal of artificial intelligence (AI) is to build autonomous agents capable of growing such open-ended repertoires of skills.

Reinforcement learning (RL) offers a mathematical framework to formalize and tackle skill learning problems. For an embodied and situated RL agent, *learning a skill* (e.g. playing chess) is about learning to act so as to maximize future *rewards* measuring progression in that skill (e.g. +1 for winning a game, -1 for losing it).[1] Extensions based on modern deep learning methods (deep RL) have recently made the headlines by solving a wealth of problems we thought only humans could solve: playing Atari video games at super-human level,[2] beating chess and go world champions,[3] controlling stratospheric balloons[4] or even maintaining plasma in fusion reactors.[5] But human chess world champions can also run, cook, draw a cat, or make a friend laugh. Humans are proficient in a wide diversity of tasks, most of which they just invent for themselves. The RL framework, in its standard form, considers a single predefined reward function and, thus, must be extended (see Figure 1, left).

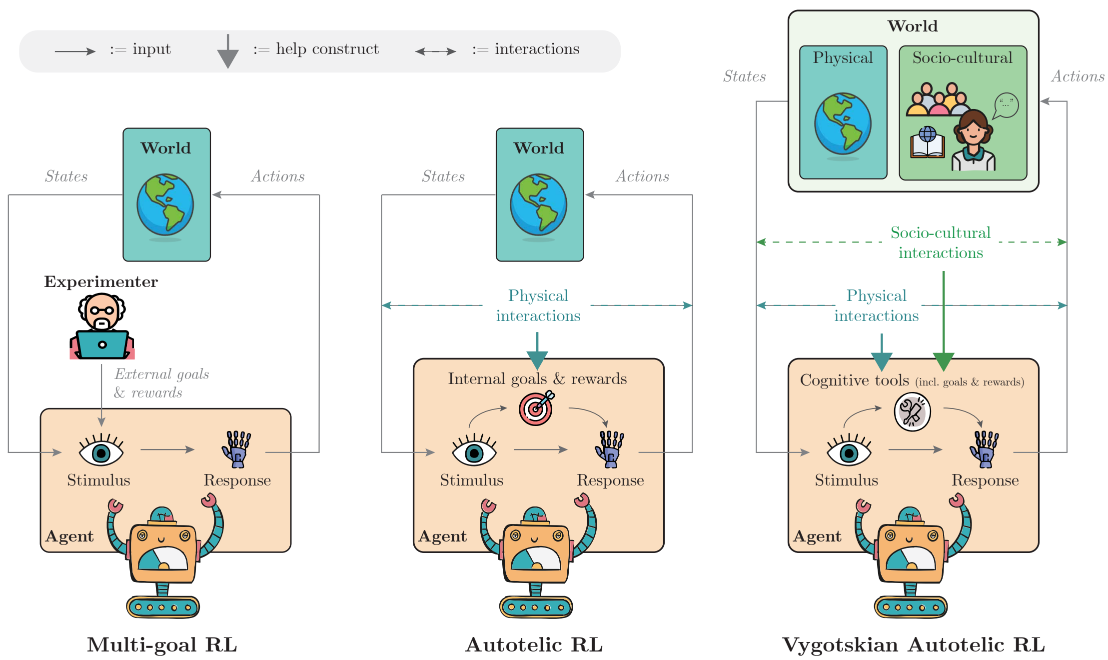

Figure 1. From multi-goal RL to autotelic RL to Vygotskian autotelic RL. RL defines an agent experiencing the state of the world as stimuli and acting on that world via actions. Multi-goal RL (left): goals and associated rewards come from pre-engineered functions and are perceived as sensory stimuli by the agent. Autotelic RL (middle): agents build internal goal representations from interactions between their intrinsic motivations and their physical experience of the world (Piagetian view). Vygotskian autotelic RL (right): agents internalize physical and socio-cultural interactions into *cognitive tools*. Here, *cognitive tools* refer to any self-generated representation that mediates stimulus and actions. This can include self-generated goals, explanations, descriptions, attentional biases, visual aids, mnemotechnic tricks, etc.

Jean Piaget, a pioneer of developmental psychology, demonstrated children's ability to set their own goals, to decide of their own learning trajectories and to develop in symbiosis with their physical environment.[6,7] His work influenced future developments in both psychology and AI. Indeed, if each skill is the association of a goal (e.g. "be a good chess player") and a policy to reach it, then *open-ended skill discovery* presupposes the ability to invent and select one's own goals and reward functions so as to progressively build repertoires of skills. *Autotelic RL* — from the Greek *auto* (self) and *telos* (goal) — extends the RL framework to build such agents.[8–15] It integrates two extensions of the standard RL framework: the ability to consider multiple goals in parallel (*multi-goal RL*) and the ability to represent and select one's own goals. Although the first extension is straightforward,[16,17] the second requires another ingredient inspired from the study of human learning: *intrinsic motivations*.

Most of human time is spent on activities that do not seem to satisfy any utilitarian end; think about children playing, or adult watching movies. Psychologists argue that such exploratory behaviors are powered by *intrinsic motivations* (IM), a set of brain processes driving us to experience situations for the mere pleasure of experiencing novelty, surprise, or learning progress.[18–22] Similar processes can be coded into artificial agents to foster spontaneous exploratory capabilities.[23–25] Among them, *knowledge-based IM* drive agents to experience parts of the environment to improve their internal models of the world and *competence-based IM* let them improve their mastery of self-generated goals.[26] In the Piagetian tradition, autotelic agents see their goal representation and goal selection mechanisms emerge from interactions between competence-based IM and their experience of the physical world (see Figure 1, middle).[8–15]

In practice, current autotelic RL implementations still lack human-like open-endedness. The goal representations emerging from their intrinsically motivated experience with the physical world end up very concrete and mostly consist in reaching target stimuli (e.g. matching their visual input with a particular target).[15] This contrasts with the wide diversity and the abstraction of goals targeted by humans. In addition, the generated goals very often belong to the distribution of previously experienced effects, which drastically limits the ability of autotelic agents to represent *creative goals*, thus to explore and undergo an open-ended discovery process.[12] Besides goal imagination, RL algorithms still lack human-like capacities in terms of generalization, skill composition, abstraction, or sample efficiency.[27,28]

The way forward might build on an alternative view of child development proposed by another pioneer of developmental psychology called Lev Vygotsky. Humans are social beings; intrinsically motivated to interact and cooperate with their peers.[29–32] For Vygotsky, linguistic social interactions such as descriptions, explanations, corrections, or play start as interpersonal processes before they are turned into *intrapersonal* cognitive processes through the process of *internalization*.[33–35] Following his vision, many psychologists,[36–38] linguists,[39–41] and philosophers[42–45] argued for the importance of socio-cultural interactions in the development of human intelligence.

What does this mean for AI? We advocate for a *Vygotskian Autotelic AI*. Specifically, we propose to immerse autotelic agents into our rich socio-cultural world; to let them interact with us and with their peers in natural language; to let them internalize these interactions and mesh them with their cognitive development (see Figure 1, right). Just like they do for humans, language and culture will help shape the agents' goal representations and generation mechanisms, thereby offering them the ability to generate more diverse and abstract goals; to imagine new goals beyond their past experience. Because they will develop at our contact, bathed in our cultures, they will learn about our cultural norms, values, customs, interests, ways of thinking; all of which would be impossible to learn in social isolation. Just like humans, machines will use language to develop higher cognitive functions like abstraction, generalization, or imagination.[38,46,47] As we will see, this process has already started.

Recent advances in AI make our proposition particularly timely. Indeed, these past years have seen a revolution in natural language processing (NLP), the set of tools designed to analyze and generate language. Generative models of language trained on gigantic amounts of text can now produce high-quality language,[48] handle multimodal inputs,[49,50] capture common sense[51] or cultural artefacts.[52,53] Pure language models like GPT-3[48] or PaLM[54] are capable of impressive zero-shot generalizations including joke explanations, arithmetic, question answering or translation and even multi-step reasoning when nudged appropriately.[55] Multi-modal variants can explain visual jokes (memes)[50] or generate impressive images from the most creative descriptions its testers can come up with.[56,57] The ongoing convergence of autotelic RL and NLP will offer a wealth of opportunities as autotelic agents learn to interact with us, learn from us and teach us back. This motivates the elaboration of a theoretical framework to understand recent linguistic RL developments and point towards future challenges in the quest for open-ended skill discovery.

This perspective extends previous arguments for a more social cognitive robotics[58–61] by proposing the integration of the autotelic RL framework with social processes supported by NLP. As a result, it will not discuss non-embodied multi-modal supervised techniques[49,50] or non-linguistic autotelic RL.[14,17,62,63] It will not cover the advances in the sub-field of RL dedicated to the computational modeling of social interactions (social RL).[64,65] Although future Vygotskian autotelic agents must incorporate social RL models, current approaches do not consider autotelic agents able to set their own goals and, as such, do not tackle the open-ended skill discovery problem, see a complete discussion in a related paper.[66]

The next section sets the background and discusses the interaction between language and thought in humans by building on the work of psychologists and philosophers (Section 1). The two following sections dive into the two key elements of a Vygotskian autotelic AI identified in Section 1: 1) the ability to exploit information contained in linguistic structure and content (syntax, vocabulary, narratives) to support the development of cognitive functions (Section 2); 2) the ability to internalize linguistic interactions within the agent to power its future autonomy and integration to the socio-cultural world (Section 3). Finally, Section 4 identifies open challenges for future research.

1 Language and Thought in Humans, a Vygotskian Perspective

Our ability to generate new ideas is the source of our incredible success in the animal kingdom. But this ability did not appear with the first *homo sapiens* 130,000 years ago. Indeed, the oldest imaginative artifacts such as figurative arts, elaborate burials or the first dwellings only date back to 70,000 years ago.[67,68] This is thought to coincide with the apparition of *recursive language*.[68–70] Which of these appeared first? Creativity or recursive language? Or did they mutually bootstrap?

Extreme views on the topic either characterize language as a pure communicative device to convey our inner thoughts (strong communicative thesis)[71,72] or, on the other hand, argue that only language can be the vehicle of our thoughts (strong cognitive thesis)[73,74] As often, the truth seem to lie in between. Animal and preverbal infants demonstrate complex cognition,[75,76] but language does impact the way we perceive,[77,78] represent concepts,[41] conduct compositional and relational thinking,[38,68,79] etc. Thus, language seems to be at least *required* to develop some of our cognitive processes (requirement thesis), and might still be the vehicle of *some* of our thoughts (constitutive thesis).[46] Interested readers can find a thorough overview of this debate in *Language and Thought* by Carruthers and Boucher.[46]

If language is required to develop some of our higher cognitive functions, then autotelic artificial agents should use it as well. But how does that work? What is so special about language? Let us start with *words*, which some called *invitations to form categories*.[77] Hearing the same word in a variety of contexts invites humans to compare situations, find similarities, differences and build symbolic representations of agents, object and their attributes. With words, the continuous world can be simplified and structured into mental entities at various levels of abstraction. The recursivity and partial compositionality of languages allow us to readily understand the meaning of sentences we never heard before by generalizing from known words and syntactic structures. On the flip side, it also supports *linguistic productivity*,[71] the ability to generate new sentences, thus new ideas, in an open-ended way. Relational structures such as comparisons and metaphors facilitate our relational thinking,[38,79] conditions our ability to compose mental images,[68] and supports our understanding of abstract concepts such as emotions, politics or scientific theories.[41,42] Finally, language is a cultural artefact inherited from previous generations and shared with others. It is the support of our cultural evolution and allows humans to efficiently transfer knowledge and practices across people and generations.[80] Through shared cultural artefacts such as narratives, we learn to share common values, customs and social norms, we learn how to navigate the world, what to attend to, how to think, what to expect from others, etc.[81] This cultural knowledge is readily accessible to children as they enter societies via social interactions and formal education. Learning language further extends our access to cultural artefacts through books, movies, or the Internet. These act as a thousand virtual social partners that we can learn from.

We now understand why language is so special. Let us focus on how it can shape cognitive development in humans and machines. Dennett, proponent of the requirement thesis, suggests that linguistic exposition alone can lead to a fundamental cognitive reorganization of the human brain.[43] He compares it to the installation of a serial virtual machine on humans' massively parallel processing brains. As a result, a slight change in our computational hardware (e.g. compared to our primates relatives) could open the possibility for any cognitive software reprogramming driven by language, in turn triggering the learning and cultural evolution of higher cognitive capacities. Carruthers, proponent of the constitutive thesis, suggests that language may have evolved as a separate module to exchange inner representations with our peers (naive physics, theory of mind, etc). This would require connections between linguistic and non-linguistic modules to allow conversions between inner representations and linguistic inputs/outputs. In the same way that humans can trigger imagined visual representations via top-down connections in their visual cortex, top-down activations of the linguistic module would create *inner speech*. This hallucinated speech, when broadcast to other modules, would implement *thinking in language*.[82] Clark advances yet another possibility, the *supra-communicative view*. Here, language does not transform the way the brain makes computation and is not the vehicle of thoughts. Instead, language complements our standard computation activities by "*re-shaping the computational spaces*," turning problems that would be out of reach to problems our pattern-matching brains can solve.[44] In that sense, language is a *cognitive tool* that enhances our cognitive abilities without altering them per se.

Vygotsky's theory brings a complementary argument to this debate. Caretakers naturally scaffold the learning experiences of children, tailoring them to their current objectives and capacities. Through encouragements, attention guidance, explanations or plan suggestions, they provide cognitive aids to children under the form of interpersonal social processes.[35] In this *zone of proximal development*, as Vygotsky coined it, children can benefit from these social interactions to achieve more than they could alone. In these moments, children *internalize* linguistic and social aids and progressively turn these interpersonal processes into intrapersonal *psychological tools*.[35] This essentially consist in building internal models of social partners such that learners can self-generate contextual guidance in the absence of an external one. Social speech is internalized into private speech (an outer speech of children for themselves) which, as it develops, becomes more goal-oriented and provides cognitive aids of the type caretakers would provide.[35,36,83] Progressively, it becomes more efficient and abbreviated, less vocalized, until it is entirely internalized by the child and becomes *inner speech*.

This section showed why language is so important and might just be required for the development of our highest cognitive functions. If we want machines to show human-like open-ended skill discovery processes, we might need to immerse them into rich socio-cultural worlds from the very beginning, just like we do with children, and equip them with tools to benefit from them. From the arguments above, we identify two key elements, see Figure 2. First, we need our autotelic agents to exploit the information contained within linguistic structures and contents. Exposed to language, agents will reorganize their internal representations for better abstraction, generalization and better alignment with human values, norms and customs (Dennett's thesis). Second, we need our autotelic agents to internalize social interactive processes, i.e. to *model social partners within themselves* (Vygotsky's internalization). Social processes, turned into intrapersonal cognitive processes will orient the agent's focus, help it decompose task or imagine goals. This inner speech generation can serve as a common currency between other modules (e.g. perception, motor control, goal generation) in line with Carruther's view and will help agent project problems onto linguistic spaces where they might be easier to solve (Clark's view). In the next two sections (2–3), we discuss these two key elements in details and reframe recent works at the intersection of RL and language in their light.

2 Exploiting Linguistic Structure and Content

In its vocabulary, syntax and narratives, language captures a lot of information about the world. According to Dennett's thesis, mere exposure to language can already help agents rewire their inner processes and develop new abilities. Recent advances in AI seem to support that idea.

Learning to abstract and generalize. Exposition to linguistic labels is known to facilitate category learning in humans,[77,78] but also in machines.[84,85] As Mirolli and Parisi defend, the repeated occurrence of a linguistic label, *red* in their example, leads to the conflation of internal

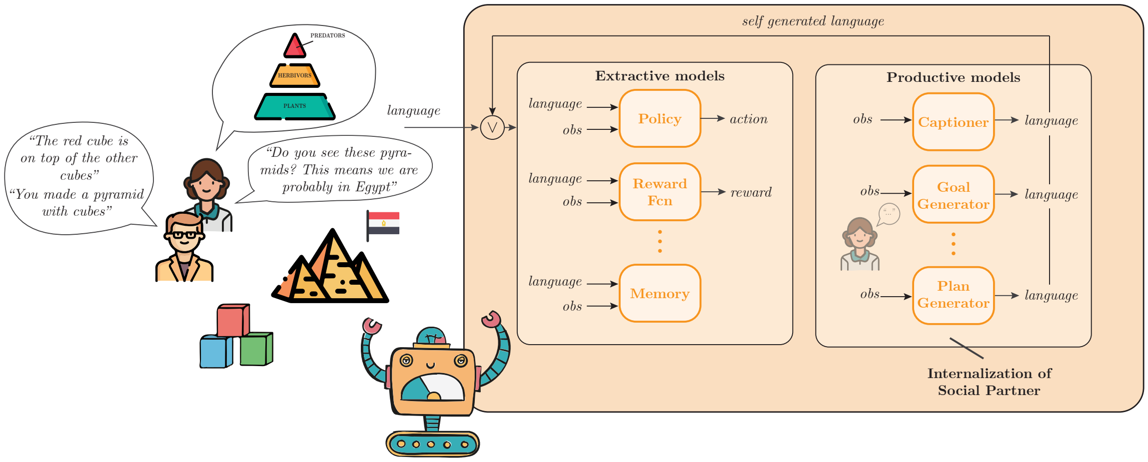

Figure 2. Socio-cultural interactions, linguistic extraction and production. Vygotskian autotelic agents are immersed into rich socio-cultural worlds where they experience a variety of linguistic feedback including descriptions, explanations or metaphors (left). They can exploit information from linguistic structures and content by conditioning their internal modules on this feedback (middle, extractive models). Finally, they learn to internalize social interactions by training productive models of language to generate feedback similar to the one they receive from others (right, productive models). This offers agents the autonomy to build their own cognitive tools, bootstrapped by socio-cultural language.

representations associated to that label (red things) which, in turn, facilitates further classifications based on the linguistic attributes.[61]

We see a similar effect in RL agents targeting linguistic goals. The exposition to aligned instructions and trajectories seem to reshape the internal representations of the agent contained within its action policy. The policy is a neural network-based function conditioned on the agent's instruction that maps the current state of the world to its next actions. By *internal representations*, we mean representations computed within the layers of the policy to facilitate the final decision making. When repeatedly asked to grasp *red objects*, the policy learns to focus on objects' colors to facilitate action selection.[12,86] *Red* is an abstraction over a continuous space of colors. It is first cultural, outside of the agent, but progressively gets internalized within the agent via a combination of linguistic exposure and decision making.

Exposed to a diversity of instructions, agents gain new cognitive abilities. The first is *abstraction*. Linguistic autotelic agents can reach and make sense of abstract goals like "*go to any circle*,"[87] relational goals "*sort objects by size*,"[88] "*put the cylinder in the drawer*,"[89] sequential goals "*open the yellow door after opening a red door*,"[90] or even learning goals "*is the ghargh edible?*."[91] Whereas handling abstract goals used to require engineers to hard-code specific goal representations and reward functions within the agent,[14,92,93] linguistic goals offer abstraction via simple linguistic interactions.[12,94] Once abstractions have been distilled within the representations of the agent, they can be leveraged to augment its exploratory capacities. Searching for novelty in a space of abstract linguistic descriptions of the world is indeed more efficient than searching for novelty in low-level sensorimotor spaces which could be trivially triggered by leaves moving in the wind or TV noise.[95,96]

A second cognitive ability is *systematic generalization*. Language-instructed agents indeed seem to demonstrate the ability to generalize to new instructions obtained by systematic recombinations of instructions they were trained on.[86] For instance, agents that learned to *grasp blue objects* and *put green objects on the table* can directly *grasp green objects* and *put blue objects on the table*.[12,86,90,94,97–101] This ability can either be encoded in learning architecture through the use of modular networks (neuro-symbolic approaches),[102] or emerge spontaneously in plain networks under the right environmental conditions.[86] Although sometimes the world does not conform to a strict linguistic compositionality, systematic generalization still supports good priors — e.g. *feeding the cat* is not a strict transposition of *feeding the plant* but they still share similarities (bringing supplies to the cat/plant).[12]

Learning to represent possible futures. After being exposed to aligned trajectories and linguistic descriptions, agents can generate concrete examples of abstract descriptions. The DECSTR approach, for example, trains a generative world model to sample from the distribution of possible future states matching a given abstract linguistic description.[13] This simple mapping supports *behavioral diversity*, the ability to represent different possible futures so as to select one to pursue. Similar setups could leverage DALL-E, an impressive text-to-image generative system.[56,57] Trained on pairs of images and compositional descriptions, DALL-E can generate high-quality images from the most twisted descriptions humans can think of. The exposition to compositional language, paired with sufficiently good learning architectures and algorithms leads to impressive visual composition abilities that could be put to use to generate visual goals, or to represent possible futures in embodied and situated agents.

Learning to decompose tasks. Vygotsky and others discovered that children's use of private speech helps them increase self-control and is instrumental to their capacity to reason and solve hard tasks[35,36,83] The ability to formulate sentences like "at the left of the blue wall," for instance, predicts spatial orientation capacities in such contexts, while interfering with adult's inner speech via speaking tasks hinders theirs.[103]

Language indeed contains cues about how to decompose tasks into sub-tasks, i.e. how to *generate good plans*. Although *gharble* is a made up word, *fry the gharble* probably involves a preparation of the gharble (e.g. peeling, cutting), some sort of oil and a frying pan.[91] *Draw an octogone* contains cues about the decomposition of the task: *octo* means 8, so we should probably do something 8 times, etc.[104] Recent AI approaches leverage these regularities by training *plan generators* from linguistic task descriptions.[88,100,104–108] Among them, Wong et al. use plan generation as an auxiliary task to train a drawing policy.[104] Generating plans to solve a particular drawing task helps shape the internal representation of the main policy which, they find, favor abstraction and generalization in the main task. Interestingly, language only shapes representations and is not required at test time, in line with the requirement thesis of Dennett.

Inspired from video games of the 80s such as *Zork*, text-based environments define purely linguistic goals, actions and states.[91,109–113] Training a policy in such environment can be seen as training a plan generator in a linguistic world model, i.e. training an inner speech to generate good task decompositions. This idea was exploited in *AlfWorld*, where a pre-trained plan generator is then deployed in a physical environment to generate sub-goals to a low-level policy.[108] Here, the abstraction capabilities of language help the plan generator solve long-horizon tasks.

The above approaches echo the thesis of Dennett (Section 1): the mere exposition to structured language, once internalized within internal modules (reward function, policy, world model) strongly shapes inner representations in new ways and supports new cognitive functions (abstraction, future states generation, compositional generalization, task decomposition, etc).

Learning from cultural artefacts. Large language models (LLM) are trained on huge quantities of text scrapped from the internet: wikipedia, forums, blogs, scientific articles, books, subtitles, etc.[48,114,115] As such, they can be seen as *cultural models* that contain information about our values, norms, customs, history, interests, etc.[52,53] This represents a great opportunity for autotelic agents to learn about us, to align with us and to better navigate our complex world. So far, only very little research has leveraged that opportunity. An example is the use of a trained LLM to act as a zero-shot planner, i.e. a plan generator.[116] Plugged with an interactive agent, the language model is used to generate sub-goals for the agent to solve a main task. Another work extracts information about complex time-extended behaviors from an LLM by asking it to score the actions available to the agent.[117] Challenge #3 of Section 4 will discuss more opportunities to harness these cultural models for open-ended skill discovery.

3 Internalization of Language Production

Agents that internalize extractive models learn to exploit information contained within linguistic vocabulary, structures and narratives. Although some of them do not require language at test time,[104] most still require an external linguistic input and, thus, cannot be considered autonomous. If we want Vygotskian autotelic agents to fully control the use of language to support cognitive functions, we must let them internalize *productive models* as well; i.e. learn internal modules that generate an artificial inner speech. This inner speech can be fed back to the extractive models such that agents fully control the whole linguistic loop (see Figure 2, right). In the literature, we found four types of productive models: trajectory captioners, plan generators, explanation generators and goal generators.

Trajectory captioners. Trajectory captioners are trained on instructive or descriptive feedback to generate valid descriptions of trajectories.[12, 118–120] In line with Vygotsky's theory, these agents internalize a model of a descriptive social partner. They generate an *inner speech* describing their ongoing behaviors just like a caretaker would. This allows them to generate new multi-modal data autonomously, and to learn from past experience via *hindsight learning*, i.e. to reinterpret their trajectory as a valid behavior to achieve the trajectory's description.[12, 88, 121–123]

Plan generators. Plan generators are both extractive and productive. Following the formalism of hierarchical RL (HRL), plan generators are implemented by a *high-level policy* generating linguistic sub-goals to a low-level policy (executioner).[124, 125] Linguistic sub-goals are a form of inner speech that facilitates decision making at lower temporal resolution by providing abstract, human-interpretable actions, which themselves favor systematic generalization for the low-level policy (see Section 2).[88, 105, 106, 108] Here again, agents internalize linguistic production to autonomously generate further guidance to themselves (task decompositions).

Explanation generators. Vygotskian autotelic agents can generate *explanations*. Using the generation of explanations as an auxiliary task was indeed shown to support causal and relational learning in complex *odd one out* tasks.[126] Note however that this approach is neither embodied, nor autotelic.

Goal generators. Some forms of creativity appear easier in linguistic spaces because swapping words, compositing new sentences, generating metaphors are all all easier in language space than in sensorimotor spaces. The IMAGINE approach leverages this idea to support *creative goal imagination*.[12] While previous methods were restricted to generate goals within the distribution of past experience (e.g. with generative models of states[10, 62, 127]), IMAGINE invents out-of-distribution goals by combining descriptions of past goals. These manipulations occur in linguistic spaces directly and are thus *linguistic thoughts* (Carruthers' view). The problem of goal imagination, difficult to solve in sensorimotor space, is projected onto the linguistic space, solved there, and projected back to sensorimotor space (Clark's view). This, in turn, powers additional cognitive abilities. First, it powers a creative exploration oriented towards objects and interactions with them. Second, it enhances systematic generalization by widening the set of goals the agent can train on.[12]

By internalizing linguistic production, IMAGINE generates goals that are both *novel* (new sentences) and *appropriate* (they respect linguistic regularities, both structures and contents).[128, 129] Social descriptions focus on objects, object attributes and interactions with these objects. Imagined goals obtained by recompositions of social ones share the same attentional and conceptual biases, e.g. by reusing semantic categories of a particular culture. Thus, cultural biases are implicitly transmitted to the agent, which forms goal representations and biases goal selection following cultural constraints.[12]

Note that productive models are very rare in the literature. In the future, Vygotskian autotelic agents must learn to internalize productive models for all types of multi-modal feedback they encounter: advice, explanations, attention guidance, motivation, instructions, descriptions, etc. It is only by learning to generate these guidance for themselves that they may gain full control on their own behavior. Furthermore, we did not encounter any work proposing autonomous corrections, augmentations or generation of cultural artefacts. In the future, agent should be able to edit cultural artefacts to build their own narrative systems, shape their own understanding of the world and others influenced by both their culture and their personal experience.

4 Open Challenges

We identify three main challenges for future research.

Challenge #1: Immersing autotelic agents in rich socio-cultural worlds. To benefit from language, Vygotskian autotelic agents must be immersed in rich socio-cultural worlds, close to ours. This will require progress along two dimensions: 1) interactivity and teachability of autotelic agents; 2) richness of their world. Sigaud and colleagues discuss the first objective through a detailed analysis of children's learning abilities and teacher-child interactions.[66] They present a checklist of properties for *teachable autotelic agents* that must serve as a roadmap towards Vygotskian autotelic agents. The second point requires to scale the richness and quantity of agent–agent and human–agent interactions. This will require the expansion of human-in-the-loop research. Either we let artificial agents enter our everyday human cultures via robotic bodies, or we let humans interact with agents in rich multi-modal synthetic worlds that must be designed for humans to enjoy. Virtual reality technology and the video game industry in general will become key elements of a Vygotskian autotelic AI.

Challenge #2: Enabling artificial mental life with systematic internalized language production Only few approaches internalize language production within agents. So far limited to few use cases, language production should concern every possible linguistic feedback agents could receive: instructions, corrections, advice, explanations, or cultural artefacts. This internal language production is akin to an artificial *inner speech*, the embryo of *artificial mental life*. Looping back to the constitutive thesis of Carruthers presented in Section 1, inner speech acts as a common currency for inner modules to exchange information. Combined with world models, inner speech could trigger the simulation of perceptual experience (images, sounds), sensorimotor trajectories, the imagination of possible futures or past memories. Observing these hallucinations, agents could produce new behaviors and new inner speech. This inner loop acts as a mental life that could help agents reason; trigger memories or mnemotechnic representations acting as cognitive aids. As noted by Dove, this account is fully compatible with the embodied hypothesis in cognitive science.[47] Following this hypothesis, thinking and modeling sensorimotor experience are one and the same. Here, language brings another set of inputs and outputs for these simulation models and the simulation of abstract content (words, analogical structures, etc) might offer us the capacity to reason abstractly.

Challenge #3: Using llms as editable and shareable cultural models. By tapping into cultural models, agents could learn about our culture. They could learn about causality, folk psychology, politeness, ethics and all these physical or cultural information that are the subject of everyday stories: fiction, news, or even simple narratives parents use to explain everyday things to children. All this information may already be captured in existing LLMs,[51,130], we may just need to learn to extract it.

Some use cases might be straightforward to implement. One can, for instance, imagine prompting pre-trained LLM with linguistic descriptions of developmental trajectories to leverage common sense knowledge and generate new goals for exploration, extracting cultural curricula. Another promising avenue is to rely on LLM knowledge to anticipate the result of agents' actions enabling them to plan in socio-cultural space. We can also envision granting agents with querying capabilities allowing them to ask for knowledge to LLM when they cannot make progress by solely interacting with their physical environment.

For now, LLM seem very difficult to steer, edit or extend. Human cultural narratives are much more malleable and can be shifted by government policies, activism, advertising, or pop culture. Future research may look into ways for agents appropriate shared cultural models contained within LLMs, to easily edit and extend them given their own experience, to share them with others.

Conclusion

This perspective builds on the autotelic RL framework inspired by the Piagetian tradition to propose a complementary view motivated by Vygotsky's work: a *Vygotskyan Autotelic AI*. First steps in that direction can already be observed in recent works at the intersection of deep RL and NLP, but a lot remains to be done. Vygotskian autotelic agents will interact with us and with our culture, exploit linguistic structures and content and finally internalize social interactions to serve as the basis of their future cognitive functions: abstraction, compositional thinking, generalization or imagination.

Vygotskian Autotelic AI opens a new research program with exciting opportunities. The current NLP revolution can be harnessed to design more interactive and teachable agents. As they interact with humans and their peers in rich socio-cultural worlds, they will learn to leverage cultural and linguistic knowledge to bootstrap their cognitive development, and may eventually contribute back to our shared cultural evolution.

Dennett used the software/hardware metaphor to describe the impact of language on our brains.[43] Current AI research mostly focuses on *hardware* by asking *how can we design better learning architectures and algorithms?* In complement, this paper suggests to focus on *software* as well: *How can we build the right socio-cultural bath which, combined with efficient hardware, will allow the emergence of human-like AI*?

References

1. Sutton, R. S. & Barto, A. G. *Introduction to Reinforcement Learning* (MIT press Cambridge, 1998). 1

2. Mnih, V. *et al.* Human-level Control Through Deep Reinforcement Learning. *Nature* (2015). 1

3. Silver, D. *et al.* Mastering the Game of Go with Deep Neural Networks and Tree Search. *Nature* (2016). 1

4. Bellemare, M. G. *et al.* Autonomous navigation of stratospheric balloons using reinforcement learning. *Nature* (2020). 2

5. Degrave, J. *et al.* Magnetic control of tokamak plasmas through deep reinforcement learning. *Nature* (2022). 2

6. Piaget, J. *The Origins of Intelligence in Children* ((Translation Margaret Cook), WW Norton & Co, 1952). 2

7. Piaget, J. *The Construction of Reality in the Child* (Routledge, 2013). 2

8. Warde-Farley, D. *et al.* Unsupervised Control Through Non-Parametric Discriminative Rewards. In *Proc. of ICLR* (2019). 2, 3

9. Eysenbach, B., Gupta, A., Ibarz, J. & Levine, S. Diversity is All You Need: Learning Skills without a Reward Function. In *Proc. of ICLR* (2019). 2, 3

10. Pong, V. *et al.* Skew-fit: State-covering self-supervised reinforcement learning. In *Proceedings of the 37th International Conference on Machine Learning, ICML 2020, 13-18 July 2020, Virtual Event*, vol. 119 of *Proceedings of Machine Learning Research* (PMLR, 2020). 2, 3, 8

11. Campero, A. *et al.* Learning with AMIGo: Adversarially Motivated Intrinsic Goals. *ArXiv - abs/2006.12122* (2020). 2, 3

12. Colas, C. *et al.* Language as a Cognitive Tool to Imagine Goals in Curiosity Driven Exploration. *Proc. NeurIPS* (2020). 2, 3, 6, 7, 8

13. Akakzia, A., Colas, C., Oudeyer, P.-Y., Chetouani, M. & Sigaud, O. Grounding Language to Autonomously-Acquired Skills Via Goal Generation. *Proc. ICLR* (2021). 2, 3, 7

14. Stooke, A. *et al.* Open-Ended Learning Leads to Generally Capable Agents. *ArXiv - abs/2107.12808* (2021). 2, 3, 6

15. Colas, C., Karch, T., Sigaud, O. & Oudeyer, P.-Y. Autotelic Agents with Intrinsically Motivated Goal-conditioned Reinforcement Learning: a Short Survey. *J. Artif. Intell. Res.* (2022). 2, 3

16. Kaelbling, L. P. Learning to Achieve Goals. *Proc. IJCAI* (1993). 2

17. Schaul, T., Horgan, D., Gregor, K. & Silver, D. Universal Value Function Approximators. *Proc. ICML* (2015). 2, 3

18. Berlyne, D. E. Curiosity and Exploration. *Science* (1966). 2

19. Gopnik, A., Meltzoff, A. N. & Kuhl, P. K. *The Scientist in the Crib: Minds, Brains, and How Children Learn.* (William Morrow & Co, 1999). 2

20. Kidd, C. & Hayden, B. Y. The Psychology and Neuroscience of Curiosity. *Neuron* (2015). 2

21. Oudeyer, P.-Y. & Smith, L. B. How Evolution May Work Through Curiosity-Driven Developmental Process. *Top. Cogn. Sci.* (2016). 2

22. Gottlieb, J. & Oudeyer, P.-Y. Towards a Neuroscience of Active Sampling and Curiosity. *Nat. Rev. Neurosci.* (2018). 2

23. Schmidhuber, J. Curious Model-Building Control Systems. *IEEE Int. Jt. Conf. on Neural Networks* (1991). 3

24. Barto, A. G. & Simsek, O. Intrinsic Motivation for Reinforcement Learning Systems. *Proc. Thirteen. Yale Work. on Adapt. Learn. Syst.* (2005). 3

25. Oudeyer, P.-Y., Kaplan, F. & Hafner, V. V. Intrinsic Motivation Systems for Autonomous Mental Development. *IEEE transactions on evolutionary computation* (2007). Publisher: IEEE. 3

26. Oudeyer, P.-Y. & Kaplan, F. What Is Intrinsic Motivation? A Typology of Computational Approaches. *Front. neurorobotics* (2009). 3

27. Witty, S. *et al.* Measuring and Characterizing Generalization in Deep Reinforcement Learning. *Appl. AI Lett.* (2021). 3

28. Shanahan, M. & Mitchell, M. Abstraction for Deep Reinforcement Learning. *Proc. IJCAI* (2022). 3

29. Tomasello, M. *The Cultural Origins of Human Cognition* (Harvard University Press, 1999). 3

30. Tomasello, M., Carpenter, M., Call, J., Behne, T. & Moll, H. Understanding and Sharing Intentions: The Origins of Cultural Cognition. *Behav. brain sciences* (2005). 3

31. Tomasello, M. *Constructing a Language* (Harvard university press, 2009). 3

32. Brewer, K., Pollock, N. & Wright, F. V. Addressing the Challenges of Collaborative Goal Setting with Children and Their Families. *Phys. & Occup. Ther. Pediatr.* (2014). 3

33. Vygotsky, L. S. Tool and Symbol in Child Development. In *Mind in Society* (Harvard University Press, 1930). 3

34. Vygotsky, L. S. Play and Its Role in the Mental Development of the Child. *Sov. Psychol.* (1933). 3

35. Vygotsky, L. S. *Thought and Language* (MIT press, 1934). 3, 5, 7

36. Berk, L. E. Why Children Talk to Themselves. *Sci. Am.* (1994). 3, 5, 7

37. Lupyan, G. What Do Words Do? Toward a Theory of Language-Augmented Thought. In *Psychology of Learning and Motivation* (Elsevier, 2012). 3

38. Gentner, D. & Hoyos, C. Analogy and Abstraction. *Top. Cogn. Sci.* (2017). 3, 4

39. Whorf, B. L. *Language, Thought, and Reality: Selected Writings of Benjamin Lee Whorf* (MIT press, 1956). 3

40. Rumelhart, D. E., Smolensky, P., McClelland, J. L. & Hinton, G. Sequential Thought Processes in Pdp Models. *Parallel distributed processing: explorations microstructures cognition* (1986). 3

41. Lakoff, G. & Johnson, M. *Metaphors We Live By* (University of Chicago press, 2008). 3, 4

42. Hesse, M. The Cognitive Claims of Metaphor. *The journal speculative philosophy* (1988). 3, 4

43. Dennett, D. C. *Consciousness Explained* (Penguin uk, 1993). 3, 4, 10

44. Clark, A. *Being There: Putting Brain, Body, and World Together Again* (MIT press, 1998). 3, 5

45. Carruthers, P. Modularity, Language, and the Flexibility of Thought. *Behav. Brain Sci.* (2002). 3

46. Carruthers, P. & Boucher, J. *Language and Thought* (Cambridge University Press, 1998). 3, 4

47. Dove, G. Language as a Disruptive Technology: Abstract Concepts, Embodiment and the Flexible Mind. *Philos. Transactions Royal Soc. B: Biol. Sci.* (2018). 3, 9

48. Brown, T. B. *et al.* Language Models are Few-Shot Learners. *Proc. NeurIPS* (2020). 3, 7

49. Radford, A. *et al.* Learning Transferable Visual Models from Natural Language Supervision. *Proc. ICML* (2021). 3

50. Alayrac, J.-B. *et al.* Flamingo: a Visual Language Model for Few-Shot Learning. *arXiv preprint arXiv:2204.14198* (2022). 3

51. West, P. *et al.* Symbolic Knowledge Distillation: from General Language Models to Commonsense Models. *ArXiv - abs/2110.07178* (2021). 3, 9

52. Hershcovich, D. *et al.* Challenges and Strategies in Cross-Cultural NLP. *ArXiv - abs/:2203.10020* (2022). 3, 7

53. Arora, A., Kaffee, L.-A. & Augenstein, I. Probing Pre-Trained Language Models for Cross-Cultural Differences in Values. *ArXiv - abs/2203.13722* (2022). 3, 7

54. Chowdhery, A. *et al.* Palm: Scaling Language Modeling with Pathways. *ArXiv - abs/2204.02311* (2022). 3

55. Creswell, A., Shanahan, M. & Higgins, I. Selection-Inference: Exploiting Large Language Models for Interpretable Logical Reasoning. *ArXiv - abs/2205.09712* (2022). 3

56. Ramesh, A. *et al.* Zero-Shot Text-to-Image Generation. *ArXiv - abs/2102.12092* (2021). 3, 7

57. Ramesh, A., Dhariwal, P., Nichol, A., Chu, C. & Chen, M. Hierarchical Text-Conditional Image Generation with Clip Latents. *ArXiv - abs/2204.06125* (2022). 3, 7

58. Dautenhahn, K., Ogden, B. & Quick, T. From Embodied to Socially Embedded Agents – Implications for Interaction-Aware Robots. *Cogn. Syst. Res.* (2002). 3

59. Zlatev, J. The Epigenesis of Meaning in Human Beings, and Possibly in Robots. *Minds Mach.* (2001). 3

60. Lindblom, J. & Ziemke, T. Social Situatedness of Natural and Artificial Intelligence: Vygotsky and Beyond. *Adapt. Behav.* (2003). 3

61. Mirolli, M. & Parisi, D. Towards a Vygotskyan Cognitive Robotics: The Role of Language as a Cognitive Tool. *New Ideas Psychol.* **29** (2011). 3, 6

62. Florensa, C., Held, D., Geng, X. & Abbeel, P. Automatic Goal Generation for Reinforcement Learning Agents. *Proc. ICML* (2018). 3, 8

63. Akakzia, A., Serris, O., Sigaud, O. & Colas, C. Help Me Explore: Minimal Social Interventions for Graph-Based Autotelic Agents. *ArXiv - abs/2202.05129* (2022). 3

64. Jaques, N. *et al.* Social Influence as Intrinsic Motivation for Multi-Agent Deep Reinforcement Learning. *Proc. ICML* (2019). 3

65. Akalin, N. & Loutfi, A. Reinforcement Learning Approaches in Social Robotics. *Sensors* (2021). 3

66. Sigaud, O., Colas, C., Akakzia, A., Chetouani, M. & Oudeyer, P.-Y. Towards Teachable Autonomous Agents. *ArXiv - abs/2105.11977* (2021). 4, 9

67. Harari, Y. N. *Sapiens: A Brief History of Humankind* (Random House, 2014). 4

68. Vyshedskiy, A. Language Evolution to Revolution: the Leap From Rich-Vocabulary Non-Recursive Communication System to Recursive Language 70,000 Years Ago Was Associated with Acquisition of a Novel Component of Imagination, Called Prefrontal Synthesis, Enabled By a Mutation that Slowed Down the Prefrontal Cortex Maturation Simultaneously in Two or More Children – the Romulus and Remus Hypothesis. *Res. Ideas Outcomes* (2019). 4

69. Goldberg, A. E. The Emergence of the Semantics of Argument Structure Constructions. In *The emergence of language* (Psychology Press, 1999). 4

70. Hoffmann, T. Construction Grammar and Creativity: Evolution, Psychology, and Cognitive Science. *Cogn. Semiot.* (2020). 4

71. Chomsky, N. *Syntactic Structures* (Mouton, 1957). 4

72. Fodor, J. A. *The Language of Thought* (Harvard university press, 1975). 4

73. Wittgenstein, L. *Philosophical Investigations* (John Wiley & Sons, 1953). 4

74. McDowell, J. *Mind and World* (Harvard University Press, 1996). 4

75. Sperber, D., Premack, D. & Premack, A. J. *Causal Cognition: A Multidisciplinary Debate* (Clarendon Press Oxford, 1995). 4

76. Allen, C. & Bekoff, M. *Species of Mind: The Philosophy and Biology of Cognitive Ethology* (MIT Press, 1999). 4

77. Waxman, S. R. & Markow, D. B. Words as Invitations to Form Categories: Evidence from 12-to 13-Month-Old Infants. *Cogn. psychology* (1995). 4, 5

78. Yoshida, H. & Smith, L. B. Sound Symbolism and Early Word Learning in Two Languages. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (2003). 4, 5

79. Gentner, D. & Loewenstein, J. *Relational Language and Relational Thought* (Erlbaum, 2002). 4

80. Henrich, J. & McElreath, R. The evolution of cultural evolution. *Evol. Anthropol. Issues, News, Rev.* (2003). 4

81. Bruner, J. *Acts of meaning.* (Harvard university press, 1990). 4

82. Carruthers, P. Thinking in language?: Evolution and a modularist possibility. In *Language and Thought* (Cambridge University Press, 1998). 5

83. Sokolov, A. *Inner Speech and Thought* (New York: Plenum Press, 1972). 5, 7

84. Schyns, P. G. A Modular Neural Network Model of Concept Acquisition. *Cogn. Sci.* (1991). 5

85. Lupyan, G. Carving Nature at Its Joints and Carving Joints into Nature: How Labels Augment Category Representations. In *Modeling Language, Cognition and Action* (World Scientific, 2005). 5

86. Hill, F. *et al.* Emergent Systematic Generalization in a Situated Agent. *ArXiv - abs/1910.00571* (2019). 6, 7

87. Janner, M., Narasimhan, K. & Barzilay, R. Representation Learning for Grounded Spatial Reasoning. *Transactions Assoc. for Comput. Linguist.* (2018). 6

88. Jiang, Y., Gu, S., Murphy, K. & Finn, C. Language as an abstraction for hierarchical deep reinforcement learning. *Proc. NeurIPS* (2019). 6, 7, 8

89. Lynch, C. & Sermanet, P. Language conditioned imitation learning over unstructured data. *ArXiv - abs/2005.07648* (2020). 6

90. Chevalier-Boisvert, M. *et al.* Baby-Ai: First Steps Towards Grounded Language Learning with a Human in the Loop. *Proc. ICLR* (2019). 6

91. Yuan, X. *et al.* Interactive Language Learning by Question Answering. In *Proc. of EMNLP* (Association for Computational Linguistics, 2019). 6, 7

92. Colas, C., Oudeyer, P., Sigaud, O., Fournier, P. & Chetouani, M. CURIOUS: intrinsically motivated modular multi-goal reinforcement learning. In Chaudhuri, K. & Salakhutdinov, R. (eds.) *Proceedings of the 36th International Conference on Machine Learning, ICML 2019, 9-15 June 2019, Long Beach, California, USA*, vol. 97 of *Proceedings of Machine Learning Research* (PMLR, 2019). 6

93. León, B. G., Shanahan, M. & Belardinelli, F. Systematic generalisation through task temporal logic and deep reinforcement learning. *ArXiv - abs/2006.08767* (2020). 6

94. Bahdanau, D. *et al.* Learning to understand goal specifications by modelling reward. In *Proc. of ICLR* (2019). 6

95. Tam, A. C. *et al.* Semantic Exploration from Language Abstractions and Pretrained Representations. *ArXiv - abs/2204.05080* (2022). 6

96. Mu, J. *et al.* Improving Intrinsic Exploration with Language Abstractions. *ArXiv - abs/2202.08938* (2022). 6

97. Hermann, K. M. *et al.* Grounded Language Learning in a Simulated 3D World. *ArXiv - abs/1706.06551* (2017). 6

98. Chaplot, D. S., Sathyendra, K. M., Pasumarthi, R. K., Rajagopal, D. & Salakhutdinov, R. Gated-Attention Architectures for Task-Oriented Language Grounding. *Proc. AAAI* (2018). 6

99. Hill, F., Mokra, S., Wong, N. & Harley, T. Human Instruction-Following with Deep Reinforcement Learning via Transfer-Learning from Text. *ArXiv - abs/2005.09382* (2020). 6

100. Sharma, P., Torralba, A. & Andreas, J. Skill Induction and Planning with Latent Language. *ArXiv - abs/2110.01517* (2021). 6, 7

101. Karch, T., Teodorescu, L., Hofmann, K., Moulin-Frier, C. & Oudeyer, P.-Y. Grounding Spatio-Temporal Language with Transformers. *Proc. NeurIPS* (2021). 6

102. Mao, J., Gan, C., Kohli, P., Tenenbaum, J. B. & Wu, J. The Neuro-Symbolic Concept Learner: Interpreting Scenes, Words, and Sentences from Natural Supervision. *ArXiv - abs/1904.12584* (2019). 6

103. Hermer-Vazquez, L. Language, Space, and the Development of Cognitive Flexibility in Humans: The Case of Two Spatial Memory Tasks. *Cognition* (2001). 7

104. Wong, C., Ellis, K., Tenenbaum, J. B. & Andreas, J. Leveraging Language to Learn Program Abstractions and Search Heuristics. *ArXiv - abs/2106.11053* (2021). 7, 8

105. Hu, H., Yarats, D., Gong, Q., Tian, Y. & Lewis, M. Hierarchical Decision Making by Generating and Following Natural Language Instructions. *Proc. NeurIPS* (2019). 7, 8

106. Chen, V., Gupta, A. & Marino, K. Ask Your Humans: Using Human Instructions to Improve Generalization in Reinforcement Learning. *Proc. ICLR* (2021). 7, 8

107. Mirchandani, S., Karamcheti, S. & Sadigh, D. ELLA: Exploration through Learned Language Abstraction. In Ranzato, M., Beygelzimer, A., Dauphin, Y., Liang, P. & Vaughan, J. W. (eds.) *Advances in Neural Information Processing Systems*, vol. 34, 29529–29540 (Curran Associates, Inc., 2021). 7

108. Shridhar, M. *et al.* ALFWorld: Aligning Text and Embodied Environments for Interactive Learning. *Proc. ICLR* (2021). 7, 8

109. Narasimhan, K., Kulkarni, T. & Barzilay, R. Language Understanding for Text-based Games using Deep Reinforcement Learning. *Proc. EMNLP* (2015). 7

110. Côté, M.-A. *et al.* TextWorld: A Learning Environment for Text-Based Games. *Comput. Games - 7th Work. @ IJCAI* (2018). 7

111. Das, A. *et al.* Embodied question answering. *Proc. CVPR* (2018). 7

112. Ammanabrolu, P. & Hausknecht, M. J. Graph constrained reinforcement learning for natural language action spaces. In *Proc. of ICLR* (2020). 7

113. Madotto, A. *et al.* Exploration based language learning for text-based games. *Proc. IJCAI* (2020). 7

114. Devlin, J., Chang, M.-W., Lee, K. & Toutanova, K. BERT: Pre-training of deep bidirectional transformers for language understanding. *Proc. NAACL* (2019). 7

115. Gao, L. *et al.* The pile: An 800GB dataset of diverse text for language modeling. *ArXiv - abs/2101.00027* (2020). 7

116. Huang, W., Abbeel, P., Pathak, D. & Mordatch, I. Language models as zero-shot planners: Extracting actionable knowledge for embodied agents. *ArXiv - abs/2201.07207* (2022). 7

117. Ahn, M. *et al.* Do As I Can, Not As I Say: Grounding Language in Robotic Affordances. *ArXiv - abs/2204.01691* (2022). 7

118. Cideron, G., Seurin, M., Strub, F. & Pietquin, O. HIGhER: Improving Instruction Following with Hindsight Generation for Experience Replay. In *IEEE Symposium Series on Computational Intelligence (SSCI)* (2020). 8

119. Zhou, L. & Small, K. Inverse Reinforcement Learning with Natural Language Goals. *ArXiv - abs/2008.06924* (2020). 8

120. Nguyen, K., Misra, D., Schapire, R., Dudík, M. & Shafto, P. Interactive learning from activity description. *Proc. ICML* (2021). 8

121. Andrychowicz, M. *et al.* Hindsight Experience Replay. *Proc. NeurIPS* (2017). 8

122. Chan, H., Wu, Y., Kiros, J., Fidler, S. & Ba, J. ACTRCE: Augmenting Experience via Teacher's Advice For Multi-Goal Reinforcement Learning. *ArXiv - abs/1902.04546* (2019). 8

123. Eysenbach, B., Geng, X., Levine, S. & Salakhutdinov, R. R. Rewriting History with Inverse RL: Hindsight Inference for Policy Improvement. *Proc. NeurIPS* (2020). 8

124. Dayan, P. & Hinton, G. E. Feudal Reinforcement Learning. In *Advances in neural information processing systems* (1993). 8

125. Sutton, R. S., Precup, D. & Singh, S. Between MDPs and Semi-MDPs: A Framework for Temporal Abstraction in Reinforcement Learning. *Artif. intelligence* (1999). Publisher: Elsevier. 8

126. Lampinen, A. K. *et al.* Tell Me Why! – Explanations Support Learning of Relational and Causal Structure. *ArXiv - abs/2112.03753* (2021). 8

127. Nair, A. *et al.* Visual Reinforcement Learning with Imagined Goals. *Proc. NeurIPS* (2018). 8

128. Runco, M. A. & Jaeger, G. J. The Standard Definition of Creativity. *Creat. Res. J.* (2012). 8

129. Simonton, D. K. Creative Productivity and Aging. In *The Wiley-Blackwell Handbook of Adulthood and Aging* (John Wiley & Sons, Ltd, 2012). 8

130. Schramowski, P., Turan, C., Andersen, N., Rothkopf, C. A. & Kersting, K. Large Pre-Trained Language Models Contain Human-Like Biases of What Is Right and Wrong to Do. *Nat. Mach. Intell.* (2022). 9

Acknowledgements

Author contributions statement

All authors reviewed the manuscript.

Additional information

The authors declare no competing interests.