*ECAI 2016*
*G.A. Kaminka et al. (Eds.)*


*doi:10.3233/978-1-61499-672-9-471*

# Is Spearman's Law of Diminishing Returns (SLODR) Meaningful for Artificial Agents?

**José Hernández-Orallo**[1]

**Abstract.** The progress of artificial intelligence is reaching a point that some research questions that were only relevant for human and other animal agents are becoming relevant for artificial agents as well. One of those questions comes from human intelligence research and is known as Spearman's Law of Diminishing Returns (SLODR). Charles Spearman, the father of factor analysis and the *g* factor (a dominant factor explaining most of the variance in cognitive tests for human populations), observed that when the analysis was restricted to the subpopulation of most able subjects, the relevance of this dominant factor diminished, as if the power of general intelligence were saturated or not fully used by the most able individuals. In about a century, there have been numerous theoretical explanations and experiments to confirm or reject Spearman's hypothesis. However, all of them have been based on human or animal populations. In this paper, we analyse for the first time whether the SLODR makes sense for artificial agents and what its role should be in the analysis of general-purpose AI. We use a synthetic scenario based on modified elementary cellular automata (ECA) where the ECA rules work as tasks and the population of agents is generated with an agent policy language. Different slices of the population by ability and of the tasks by difficulty are analysed, showing that SLODR does not really appear. Indeed, even if very slightly, we find the reverse, i.e., that more correlation takes place for more able subpopulations, what we conjecture as the Universal Law of Augmenting Returns (ULOAR).

## 1 INTRODUCTION

While more and more specific applications are being successfully solved by AI systems, the field is also progressing in the development of systems that are able to solve a wider range of problems, usually after a long training. Reinforcement learning [49], cognitive developmental robotics [6] and machine learning in general incarnate —or are integrated into— autonomous agents that solve a range of tasks. Some recent developments have displayed significant performance in a variety of tasks, at least in some domains. For instance, [39] combine reinforcement learning and deep learning to attempt a diverse set of Atari 2600 videogames. Although the system has to relearn from scratch when the game changes slightly, it is still a general-purpose technique, which can be evaluated for a range of tasks.

Despite the great number of competitions in AI for particular applications, some competitions and benchmarks are also moving in the direction of more general-purpose systems (see [23]for a full account). These include the general game playing AAAI Competition [17, 18], the reinforcement learning competition [54, 10] (including, e.g., the 'polyathlon', with several domains), the genetic programming benchmarks [38, 53], the general video game competition [43, 42], and the arcade learning environment [2, 43] (including the Atari 2600 videogames mentioned above).

However, there are some questions that arise when one considers a wide range of tasks, both when designing systems to behave well for them or when designing tests to evaluate these systems. This discussion is especially controversial when one wants to consider *all possible tasks*. On one hand, if one considers every possible problem's output as equally likely (technically known as "block uniformity" [29], with the uniform distribution being a special case) then we have the conditions for the so-called no-free-lunch theorems [57, 56], leading to the conclusion that, on average, no method can be better than any other. According to this, a general-purpose system and, indeed, the very concept of 'general intelligence' would be impossible [14]. On the other hand, if one considers problems as programs, then a uniform distribution is not possible. Instead, any universal distribution can be assumed, which leads to the theory of universal prediction using algorithmic probability developed by Solomonoff in the 1960s [44, 45]. This has influenced several approaches based on algorithmic information theory about how tasks can be generated and weighted in definitions of intelligence [11, 19, 33, 25, 13, 30] and how theoretically general agents can be defined [28], only if weakly optimal or suboptimal in general [41, 34]. Nevertheless, the idea of general intelligence makes sense theoretically in this context: some agents can be better than others in general.

The experimental and theoretical analysis of AI agents that are devised and evaluated for a range of tasks has led to an approaching to some similar ideas from the area of human intelligence evaluation. The use of IQ tests for the evaluation of AI systems has been advocated for by some [5, 4] but it has been criticised by others for being anthropocentric [12, 27]). But other concepts and tools from psychometrics, such as item response theory and the use of task difficulty to analyse the landscape of problems, are being vindicated in artificial intelligence as well, under the term universal psychometrics [26, 24]. In fact, one of the problems of the use of a universal distribution of tasks for defining the general problem of intelligence can be addressed differently if one considers a uniform distribution of difficulties, a uniform distribution of policies per difficulty with finally leaving the universal distribution to the conditional probability of a task given an acceptable policy [21]. This replaces the notion of a task-general intelligence (the so-called universal intelligence) as addressing a diversity of tasks to that of a policy-general addressing a diversity of solutions, expressed under a policy description language.

This debate replicates the controversy in psychometrics between the IQ scores (results of the IQ tests, which depend on the task distribution used in a test) and the *g* scores (a magnitude derived from the estimated value of a latent factor, the *g* factor, which is more independent to the particular task distribution used in a test). The *g* factor

[1] DSIC, Universitat Politècnica de València, email: jorallo@dsic.upv.es

derives from the so-called positive manifold, one of the most replicated experimental findings in the analysis of human intelligence. The positive manifold indicates that given any test composed of a set of (abstract) cognitive tasks we will find a high correlation in the results produced by a human population. In other words, those who perform well on some tasks will usually perform well on any other. This supports the idea of general intelligence.

For artificial intelligence, this suggests the following question: if we aim at building more general AI systems, will it be the case that those that are better for some tasks will also be better for other tasks? Is that a necessary or a contingent *property*? In order to study this question, however, there is an important consideration: two equally-general systems may have different levels of ability (or general intelligence). It is for those that are more intelligent we expect this property to hold stronger. To put an extreme negative case, a random agent is completely general, but not intelligent. We do not expect this property to hold for random agents, despite their 'generality'. This suggests that the positive manifold will only start to be observed for artificial agents when we have a population of minimally intelligent agents. As long as AI progresses towards more generally intelligent agents, this positive manifold would start to appear and then become stronger. Surprisingly, in the realm of human intelligence, Spearman found exactly the reverse observation. By taking subpopulations of more able humans, the positive manifold was weaker, something that was later known as Spearman's Law of Diminishing Returns (SLODR).

In this paper, we introduce a simple, but effective, setting to analyse these questions for artificial agents. We adapt a class of tasks consistent of elementary cellular automata (ECA) where we have introduced an agent that interacts within these worlds (the ECA rules), as done in [20]. Using this simple world and an elementary agent language, we can analyse *all tasks* and all possible policies (solutions) up to a certain size (determining the difficulty of the policy), so really having a diversity of solutions to analyse whether some degree of positive manifold appears. More interestingly, we can easily analyse different subsets according to their average performance on all policies (or slices of appropriate difficulty) and study whether the SLODR holds or not.

The rest of the paper is organised as follows. Section 2 reviews the notions of positive manifold, *g* factor and Spearman's Law of Diminishing Returns (SLODR) and some of the explanations and experiments performed to support or reject the law. Section 3 introduces the environments and agents used for the analysis, describing how they work and showing a few examples. Section 4 performs two different experiments with the goal of discovering whether the SLODR holds or not. Section 5 discusses the results and its implications. Finally, Section 6 closes the paper with new questions and future work.

## 2 SPEARMAN'S LAW OF DIMINISHING RETURNS (SLODR)

There are many kinds of cognitive tests that can be applied to humans. Some of them compose the popular IQ tests, whose development started about a century ago. Charles Spearman was one of the pioneers of a numerical analysis of human intelligence, by compiling the results of several tests on human populations. He started to use the recently introduced notion of correlation to analyse the results. He found one important phenomenon: when he analysed a set of different tests taken by the same population, he found a positive average correlation in their results. In other words, the individuals that obtained good results for some tests usually obtained good results for the rest. This correlation was stronger the more culture-fair and abstract the tests were. This phenomenon was known as the 'positive manifold' [46, 47].

It is important to clarify that this phenomenon is not a property of the tests alone nor a property of the population alone. A correlation is clearly an effect that takes place for two subjects for a set of tests, but the average correlation is calculated from the correlation matrix, thereby involving both the population and the tests. Nevertheless, the positive manifold appeared again and again for different human populations and different sets of tests, provided they were not too linked to particular cultural or educational backgrounds (e.g., a chess-playing test and a Korean vocabulary test). Spearman tried to understand the findings through the invention of a rudimentary factor analysis. He identified a dominant *latent factor* that explained much of the variance, and called it the *g* factor. Since then, this factor has been one of the most relevant (and replicated) findings in psychometrics [31, 48] and has been found to predict many facets of life, from academic performance to (lack of) religiosity in humans.

The dominance of *g* and its explanatory character for the positive manifold led to the association of *g* with general intelligence, a latent factor that pervaded or dominated all other factors and facets of intelligence. Of course, this interpretation has been challenged many times, even if *g* appears again and again.

Still more controversial than the interpretation of the *g* factor is another finding that Spearman discovered. He calculated the strength of *g* on subpopulations of different abilities. In particular, in one of the analysis, he separated the results of several tests on a human population into two groups, group $A$ with normal abilities and $B$ with low abilities. After the split, he analysed the correlation matrices separately. The result was that the mean correlations for group $A$ were 0.47 but the mean correlations for group $B$ was 0.78. Note that this does not mean that group $A$ had worse results (in fact, it was precisely the group with highest average results), but rather that the *proportion of the variance* explained by *g* for the low-ability group was much higher than for the normal-ability group. This result was striking, especially if *g* is understood as general intelligence. It looked as if the more intelligent a population is, the less important *g* would be, in relative terms, to explain its variability. This observation turned to be known as Spearman's Law of Diminishing Returns (SLODR). The finding was replicated many times since then with different experimental settings [9, 8, 50]).

Spearman looked for an explanation and found it in the *law of diminishing returns* in economics. Many processes that are affected by many factors do not grow continuously as the result of the increase of one factor, so the influence of a single, albeit dominant, factor can become less relevant at a given point, being saturated. Spearman expressed it in this way: "the more 'energy' a person has available already, the less advantage accrues to his ability from further increments of it" [47, p. 219].

But this simile was not an explanation. Spearman postulated the "ability level differentiation", which considered that challenging items (those that can only solve the more able individuals) require the combination of many skills, and the prevalence of *g* would be slower. Basically, for the easy items, the general intelligence or some general resources would be the only available skills for low-ability subpopulations. Detterman and Daniel [9] argued similarly that if "central processes are deficient, they limit the efficiency of all other processes in the system. So all processes in subjects with deficits tend to operate at the same uniform level. However, subjects without deficits show much more variability across processes because they do not have deficits in important central processes". Other explana-

tions were introduced, such as that the "genetic contribution is higher at low-ability levels" [8] .

On the other hand, not only the above explanations but the experimental evidence itself have been contested. One common counter-explanation of the phenomenon argues that it is not that *g* is less important for able subjects, but that they find many of the problems in the tests less challenging than the normal population and then they are not forced to use general intelligence as they can solve the problems without deep thinking, i.e., more mechanically. In other words, the use of the same tests for both groups would be creating the effect. In fact, Fogarty and Stankov [16] performed an experiment where the more able group had to solve problems of higher difficulty whereas the less able group had to solve problems of lower difficulty. Under these conditions SLODR did not only appear but even the more able group showed higher correlations! This seems to agree with the idea that general intelligence is used when the individual finds a problem challenging. It is important, however, to check that the difficult problems are created without the use of spurious complications, in order to prevent that more difficult items are more specialised than the simple items. For instance, in number series problems, one can create a complex series by using the Fibonacci series. This, however, will just assess whether the subject has some particular mathematical knowledge, not really expecting that the subject is going to discover the Fibonacci series from scratch. This was already warned by Jensen, pointing out "that it is the highly g-loaded tests that differ the least in their loadings across different levels of ability, whereas the less g-loaded tests differ the most" [32]. Usually, problems featuring abstract thinking (inductive inference, analogies, etc.) are those with higher *g* loadings.

Nevertheless, one of the most relevant criticisms (or explanations), which will reappear later on in this paper, had a more statistical character. Jensen [31, p. 587] argued that the subgroups with higher abilities had lower variance than the subgroups with lower variance. This may be caused by the way the tests are designed to cover a wide range of subjects or the way the two groups are split, but the different variances were generally the case. As a consequence, the relative relevance of *g* would be lower for more able groups as there is less variance to explain.

All of the above suggests that there are several methodological problems about the analysis of SLODR in human intelligence, starting from putting into question all results for which both groups do not have the same variance and also those that include spurious problems or sample the populations in ways to get the same variance by introducing some other confounding factors. In the end, Murray et al. argue that SLODR could just be "a statistical artifact" [40].

In what follows we take a different perspective of the debate by using artificial tasks and artificial subjects. This can help us to rule out some of the confounding factors by focussing on a controlled experiment, where we can play with the population of agents and the choice of tasks more freely. Nevertheless, our interest is to analyse whether SLODR happens or not for artificial agents, and see whether the results can tell us something about the construction and evaluation of general-purpose AI agents.

# 3 A SETTING FOR ARTIFICIAL TASKS AND AGENTS

In this section, we are going to adapt the simple setting introduced in [20]. This is an appropriate scenario for practical reasons. First, it is more illustrative to use minimalistic environments where the number of observations and actions are extremely reduced, while still having some relatively rich phenomena with very simple transition functions. Second, we are interested in simplistic policy languages in order to be able to evaluate a large amount of agents quickly.

## 3.1 Agent-populated elementary cellular automata: definition and examples

The environments we will work with are composed of an elementary cellular automaton (ECA) [55] for the space $\mathcal{S}$ and the transition function $\tau$, but we will let an agent see and modify part of the usual behaviour of the automaton. The following definition specifies the complete behaviour of this kind of environment:

**Definition 1** *A single-agent elementary cellular automaton (SAECA) is a tuple* $\langle S, \tau, \rho, \pi, \vec{\sigma}^0, \nu, p^0 \rangle$. *The state space* $\mathcal{S}$ *is represented by a one-dimensional array of bits or cells* $\vec{\sigma} = \langle \sigma_1, \sigma_2, \ldots, \sigma_m \rangle$, *also known as* configuration. *We consider the array to be finite (*$|\vec{\sigma}| = m$*) but circular in terms of neighbourhood (*$\sigma_0 = \sigma_m$ *and* $\sigma_{m+1} = \sigma_1$*). There is an initial array* $\vec{\sigma}^0$, *also known as* seed. *The transition function* $\tau$ *is given by a number* $\nu$, *as any of the* $2^{2^3} = 256$ *rules that can be defined looking at each cell and its two neighbours according to the numbering scheme convention introduced in [55]. For instance, the following transitions for each triplet define an ECA rule:*

| *111* | *110* | *101* | *100* | *011* | *010* | *001* | *000* |
|---|---|---|---|---|---|---|---|
| *0* | *1* | *1* | *0* | *1* | *1* | *0* | *1* |

*The digits of the second row represent the new state for the middle cell after each transition, depending on the triplet. In the above case, 01101101, in binary, corresponds to decimal number 109, the ECA rule number with Wolfram's convention. Given this rule, the array 01100 would evolve in the following way, looping at the end:*

*01100*
*01101*
*01111*
*11001*
*01001* ↰
*11001*

*Given the behaviour of the space, we consider just one agent* $\pi$. *The agent is located at one cell (its position* $p$*) with* $1 \leqslant p \leqslant m$, *which is initially* $p^0$. *The set of observations* $\mathcal{O}$ *is given by two bits* $\langle \sigma_{p-1}, \sigma_{p+1} \rangle$ *representing the contents of the left and right neighbouring cells respectively, i.e.,* $\sigma_{p-1}$ *and* $\sigma_{p+1}$. *The actions* $\mathcal{A}$ *are given by a 'move' and an 'upshot', denoted by the pair* $\langle V, U \rangle$. *The ordered set of moves is given by* { `left`*=0,* `stay`*=1,* `right`*=2* }, *and the ordered set of upshots is* { `keep`*=0,* `swap`*=1,* `set0`*=2,* `set1`*=3* }, *which respectively mean that the content of the cell where the agent is does not change, the content of the cell is swapped (*$0 \to 1$, $1 \to 0$*), the content is set to 0 and the content is set to 1. The rewards are calculated in the following way. If the agent is at position* $p$ *at time* $t$, *then we use this formula:*

$$r^t \leftarrow \sum_{j=1..\lfloor m/2 \rfloor} \frac{\sigma^t_{p+j} + \sigma^t_{p-j}}{2^{j+1}}$$

*which counts the number of 1s which are in the neighbourhood of the agent, weighted by their proximity. It is easy to see that* $0 \leqslant r^t \leqslant 1$. *Basically, the goal of the agent is to be surrounded by the highest number of 1s possible, by creating them or by exploiting the changes performed by the ECA rule.*

*The order of events for each step in the system is: observations are produced, actions are performed, the automaton is updated and finally, rewards are produced.*

Note that the environment is parametrised by the original contents of the array $\sigma^0$, the ECA rule number $\nu$, and the original position of the agent $p^0$. Given an environment and a computable agent, the evolution of the system is computable and deterministic.

Let us see a few examples of how these environments work. Figure 1 shows the evolution of several environments with seed "01010101010101010101", and several values of $\nu$. We do not include any agent in the trials in this first figure. As a result, the space-time diagram after 200 iterations is the same as a classical elementary cellular automaton with each number $\nu$ (see, e.g., [55]).

### 3.2 Including agents: an agent policy language

Let us now explore what happens when we include agents in these environments. We new a language for expressing the agents. There are a few agent languages in the literature (see, e.g., [3, 35, 1]), but they are too oriented towards the architecture, are too focussed on Markov Decision Processes or are not sufficiently minimalistic for bounding their size and having some interesting programs. We present a very minimalist language, also taking into account the minimalist environment.

**Definition 2** *The agent policy language* APL *is given by a memory (or history) binary array* $mem$*, initially empty (and not circular), and an ordered set of instructions* $\mathcal{I}$ = { `back`*=0,* `fwd`*=1,* `Vaddm`*=2,* `Vadd1`*=3,* `Uaddm`*=4,* `Uadd1`*=5* }*. The numbers on the right will be used as shorthand for the instruction. For instance, the string 22142335 represents a program in* APL*. A program or policy* $\pi$ *is a sequence of instructions* $\iota_1, \iota_2, ..., \iota_{|mem|}$ *in* $\mathcal{I}$*. The interpreter works on its memory by using two accumulators* $V$ *and* $U$*, and the action is given by the result of the accumulators at the end of the process. Namely:*

1. *Read the observation* $\langle \sigma_{p-1}, \sigma_{p+1} \rangle$ *and its elements being appended to the history array* $mem$*.*
2. *Place the memory pointer* $b$ *at the end of* $mem$*.*
3. $V \leftarrow$ `stay`
4. $U \leftarrow$ `keep`
5. *forall* $\iota \in \pi$
6. *case* $\iota$*:*
7. `back` : $b \leftarrow \max(b-1, 1)$
8. `fwd` : $b \leftarrow \min(b+1, |mem|)$
9. `Vaddm` : $V \leftarrow (V + mem_b)$ *mod* 3
10. `Vadd1` : $V \leftarrow (V + 1)$ *mod* 3
11. `Uaddm` : $U \leftarrow (U + mem_b)$ *mod* 4
12. `Uadd1` : $U \leftarrow (U + 1)$ *mod* 4
13. *end case*
14. *endfor*
15. *return* $\langle V, U \rangle$

Let us see an example. If an agent is located at the fifth position of the configuration 000110111 and has a current history $mem$ = 111010 then the observations 1 and 0 will be appended to $mem$, leading to $mem$ = 11101010. If the policy 20242335 is applied, we start with $b = 8$, $V = 0 =$ `stay` and $U = 0 =$ `keep`, and we have the following execution:

1. $\iota_1 = 2 =$ `Vaddm`, $V \leftarrow (V + mem_8) \bmod 3 = 1 =$ `stay`.
2. $\iota_2 = 0 =$ `back`, $b \leftarrow max(8-1, 1) = 7$.
3. $\iota_3 = 2 =$ `Vaddm`, $V \leftarrow (V + mem_7) \bmod 3 = 2 =$ `right`.
4. $\iota_4 = 4 =$ `Uaddm`, $U \leftarrow (U + mem_7) \bmod 4 = 1 =$ `swap`.
5. $\iota_5 = 2 =$ `Vaddm`, $V \leftarrow (V + mem_7) \bmod 3 = 0 =$ `left`.
6. $\iota_6 = 3 =$ `Vadd1`, $V \leftarrow (V + 1) \bmod 3 = 1 =$ `stay`.
7. $\iota_7 = 3 =$ `Vadd1`, $V \leftarrow (V + 1) \bmod 3 = 2 =$ `right`.
8. $\iota_8 = 5 =$ `Uadd1`, $U \leftarrow (U + 1) \bmod 4 = 2 =$ `set0`.

After this program, which is run internally, we obtain the action that the agent will perform on the environment, which is given by $\langle V, U \rangle = \langle 2, 2 \rangle = \langle$ `right`, `set0` $\rangle$. This means that the agent will move right and set the content of the cell to 0.

While the class of policies generated by this language is infinite, the language is still not universal, and all (finite) programs end. The goal of this language is to be able to express some simple policies that may be useful in the environment.

Figure 2 shows how the environment with elementary cellular automaton number 110 varies for several agent policies. The resulting space-time diagram patterns are different. Similar things (where differences are more visible with respect to the corresponding diagram in Figure 1) happen with rule number 164 (Figure 3).

We define $\mathbb{R}$ as the (expected) response (the result) of agent $\pi$ in task $\mu$, which is calculated as an average of the rewards $r^t$ for the 200 steps $t$. For instance, in Figure 3 policy 23555 for rule 164 seems to have higher $\mathbb{R}$ than policy 24 for the same rule.

After introducing the environments (tasks) and agents (policies), in the following section we explore SLODR using subpopulations.

## 4 ANALYSIS OF SUBPOPULATIONS

Using the agent policy language APL defined above we generated 400 agents with their instructions chosen uniformly from the instruction set and a program length also uniformly distributed between 1 and 20. We evaluated each agent with all the 256 possible ECA rules, with 21 cells, fixed initialisation (seed) of "01010101010101010101", using 100 of iterations per trial.

### 4.1 Experiment 1: confounding factors

From the $256 \times 400$ results, we scaled them task per task so that for each task (ECA rule) we had mean 0 and standard deviation 1. As we will work with Pearson (linear) correlations, this scaling does not affect the correlations, but allows a better aggregation to determine the abilities of each agent. Also from the results, we calculated the $256 \times 256$ correlation matrix for the 256 rules. From all the correlations ($\frac{256 \times 255}{2} = 32640$), 29612 were positive. The average correlation was 0.146. Then we averaged the results for each agent to get their average score. We sorted agents per score and split the agent population according to different quantiles (from best to worst). This is shown in Figure 4.

Different size of the bins (subpopulations) were used for the quantiles. On the left figure, the black cross on top represents one bin with the whole population (400 agents), with an average correlation of 0.146, as said above. The second shape (using orange triangles) is formed by the 51 possible bins using agents 1..350, 2..351, ..., 51..400. For smaller bins underneath we see that the average correlation decreases (in other colours). If we look at the concave shapes we clearly see that the average correlation is not the same for the whole range, with smaller values for middle quantiles (around 0.5 on the $x$-axis). In fact, we see correlations are higher for the more able group (high performance, lower quantiles) and the less able group (low performance, higher quantiles).

Trying to interpret these first results, we can recall Jensen's criticism in section 2. What we observe can be easily explained by the choice of best or worst subsamples. These have a tendency to agree

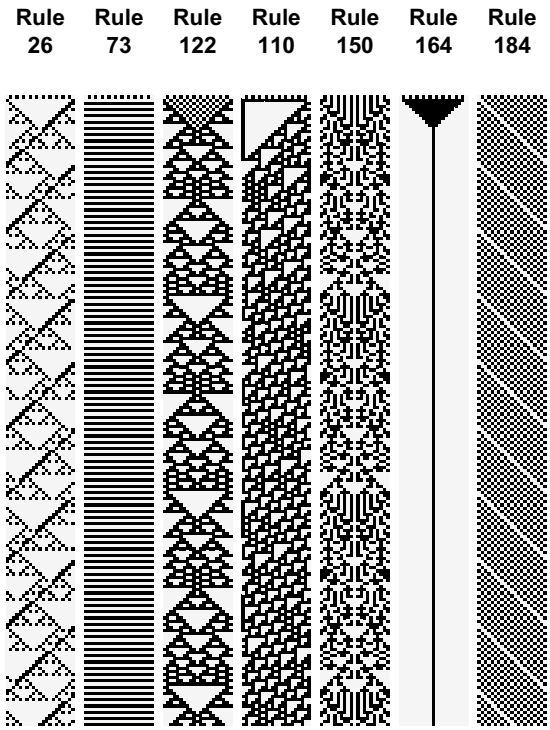


**Figure 1.** Space-time diagram (evolution for $t = 200$ steps) of several elementary cellular automata without agent. The initial array (seed) is always 01010101010101010101, whose length is 21 bits.

on more tasks and should not be interpreted as any common factor. In fact, the right plot shows the variance for each bin, which explains most of what happens on the left plot.

Nevertheless, this first experiment still shows several things. Using some quantiles split by performance, there are important differences. Of course this is not supporting any law of diminishing returns for the middle quantiles, but just a consequence of the different variances, as argued by Jensen. Also interesting is the fact that we get some positive, albeit small, average correlations, even if we are using randomly-generated agents for all possible tasks (all ECA rules). This is given by the reward mechanism, which is the same for all tasks (having 1s in the surrounding cells) and there are some agents that go well for this reward criterion disregarding the task.

In order to analyse the relevance of the reward criterion, we perform a second experiment where the reward mechanism is being mirrored half of the times (so agents cannot specialise to it). By mirroring we mean changing the sign of the reward, so now the goal is to be surrounded by as many 0s as possible. Also, the agent policy language is modified so that agents can now see the rewards. These small changes lead to very important changes in Figure 5 (left), where we now used 256 agents instead of 400. The top black cross uses all tasks (256) and all agents (256) together. The second shape (orange triangles) shows 17 bins, using agents $1..240$, $2..241$, $\ldots$, $17..256$, and so on for the other shapes, according to the sizes shown in the legend.

## 4.2 Experiment 2: variance and difficulty

The average correlation almost disappears. It is now just 0.004. Again, Figure 5 (left) slices the agents in bins by their average abilities and we have shapes that are similar to the previous experiment.

However, we now do an extra change in our analysis. Figure 5 (right) also slices the tasks by *difficulty*. We evaluate the more able agents with more difficult tasks. In order to do this, we calculate difficulty of a task following [22], where we simplify the estimation of difficulty here by only considering the length of the policies (and not the execution steps as all policies have a finite execution time):

$$\hbar^{[\epsilon]}(\mu) \triangleq \min_{\pi \in \mathcal{A}^{[\epsilon]}(\mu)} L(\pi) \tag{1}$$

i.e., the difficulty of a task $\mu$ is the length of shortest policy $\pi$ that is acceptable for the task. Note that this is not the Kolmogorov complexity of the task (i.e., the shortest description for the task) but rather the shortest description of any (acceptable) solution for the task. Acceptability is defined using a tolerance $\epsilon$:

$$\mathcal{A}^{[\epsilon]}(\mu) \triangleq \{\pi \,:\, \mathbb{R}(\pi, \mu) \geqslant 1 - \epsilon\} \tag{2}$$

i.e., the set of all acceptable policies for a task $\mu$ is given by those policies whose expected response is above a threshold, given by the tolerance $\epsilon$. Recall that we defined expected response $\mathbb{R}$ as the average reward result of agent $\pi$ in task $\mu$.

Given this approximation to difficulty, we chose tolerance to be the response that separates the 10% best agents for each task and we sliced tasks by difficulty, using the same bin size than for the agents. As we only generated 256 agents in this experiment, the sizes where also the same. In summary, Figure 5 (right) shows different shapes. As mentioned above, the top black cross uses all tasks (256) and all agents (256) together, with 0.004 correlation, and the second shape (orange triangles) shows 17 bins, using agents $1..240$, $2..241$, $\ldots$, $17..256$, and so on. But now, for each of the bins, we also slice the problems (tasks) according to their difficulty. For instance, for the first bin of the orange triangles, the most able agents $1..204$, we calculate the correlation with only the most difficult tasks $1..204$.

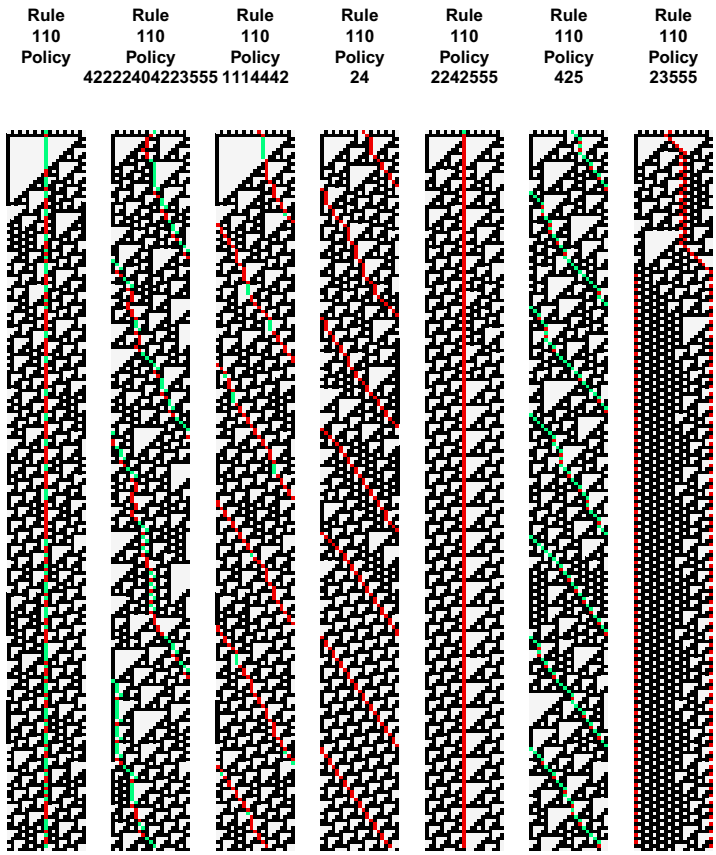


**Figure 2.** Space-time diagram (evolution for $t = 200$ steps) with different agent policies for the elementary cellular automaton with rule 110. The initial array (seed) is always 01010101010101010101, whose length is 21 bits. The agent is represented by a red dot when the cell has a 0 (like the white ones) and by a green dot when the cell has a 1 (like the black ones). The leftmost diagram is the empty policy ($\langle$`stay`, `keep`$\rangle$).

The slicing by ability and corresponding difficulty for each group now shows a very different picture. Figure 5 (right) shows some slope in the distribution of results, where we find higher correlations for higher abilities (lower quantiles). This is exactly the reverse of Spearman's Law of Diminishing Returns.

## 5 DISCUSSION

Before jumping into any conclusion, let us first analyse the results of this particular experiment. We are generating agents with a very simple policy language. Still, it can now access the rewards and compute actions with them so that meaningful policies are generated. For instance, the policy that repeats the previous action if the reward is good and do another action otherwise can be coded with a relatively short program in this language. Nevertheless, we cannot expect any agent that is especially good. Accordingly, many agents are completely lost in the environments. However, it is precisely a basic scenario we wanted to explore first, resembling some kind of minimal artificial life situation where we can consider all agents up to a certain complexity and see if any correlation appears, even if small. The simplicity of the policies was also useful for a second, very important thing. Difficulties are estimated from first principles also using the agent policy language. As seen in eq. 1, difficulty is calculated as the length of the shortest acceptable policy. This can only be estimated in a reasonable amount of time with standard hardware if programs do not get very large. Of course, the simple scenario leads to very small correlations, since we have very simple agents, and even very small correlations (once the rewards were mirrored), but the results are consistent to a low expectation about the abilities of these agents. This low correlation is also consistent to the very intuition under ULOAR.

However, the interesting point is that we can study task correlation in a very controlled experiment and find some trends by slicing per ability and difficulty. If we focus on the more able agents, it is not that they are just better for a random sample of tasks, but that they have a slight higher chance of getting more difficult problems right. The positive manifold starts appearing, the embryo of some kind of general ability may be appearing here. This suggests the hypothesis that given a population of agents, the more generally intelligent and diverse they are, the stronger the positive manifold will be. We can call this hypothesis the **universal law of augmenting returns** (ULOAR) as the opposite to SLODR. In fact, as argued before, the ULOAR makes more sense for AI, there is no reason to think that for artificial agents we may find some kind of saturation, once the tendency is initially found at very low degrees of general ability.

Of course, we cannot extrapolate from a single experiment that just shows a slightly higher (yet very small) correlation for the more able groups, but this contributes to the intuition that, for artificial agents, SLODR may not hold in general. Also, we cannot extrapolate this for human populations, and it is still unclear whether SLODR holds for humans or, more precisely, whether it holds for some human populations with some distributions of tests. But when one conceives artificial agents of a wide range of resources and algorithms, the existence of SLODR looks very counterintuitive in hindsight.

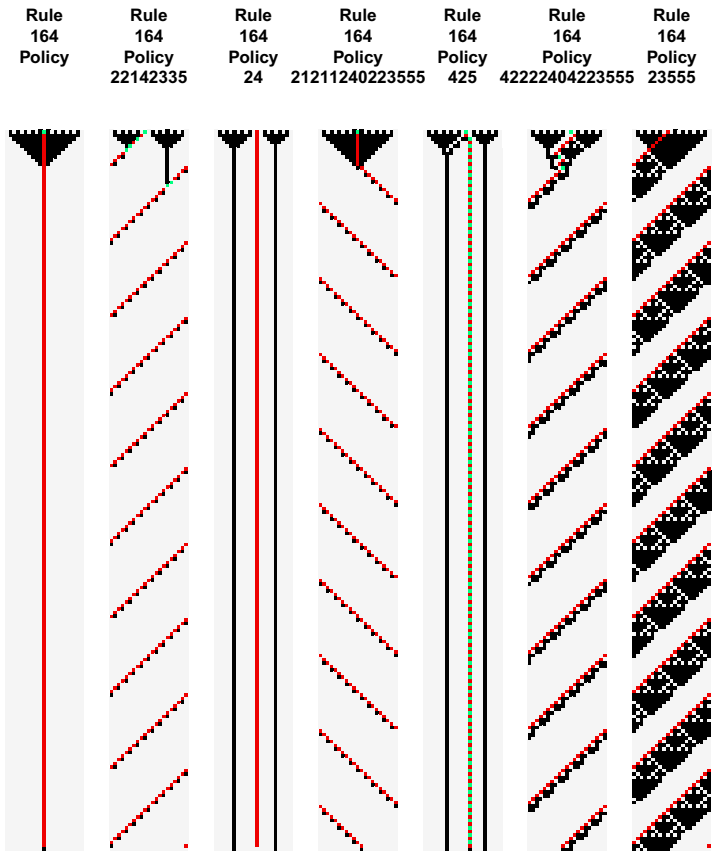


**Figure 3.** Same as Figure 2, with rule 164 and other policies.

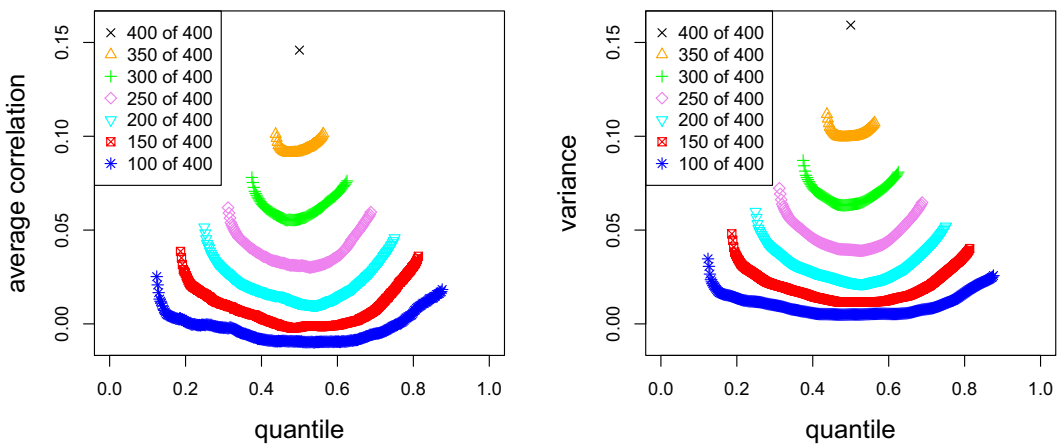


**Figure 4.** Average correlations for 400 agents (randomly-generated programs) and 256 tasks (all the ECA rules). Left: average correlation per quantiles using several bin sizes where agents are sorted by overall performance. Right: the variance of the bins.

## 6 CONCLUSION

In this paper we have argued that some questions that have been relevant for human intelligence may become soon important for artificial intelligence as well. One of these questions is the existence of general intelligence and how it can be measured and distinguished from the performance in particular skills. Given the notion of general intelligence as performance in a range of tasks, we have followed the recent theoretical and experimental analysis of the problem in AI (from the no-free-lunch theorems to algorithmic information theory) and focussed on one particular phenomenon found in human populations, known as Spearman's Law of Diminishing Returns: the positive manifold (the positive correlation of results for a set of cognitive tasks) has been shown to be stronger for less able subpopulations than more able subpopulations.

The choice of this phenomenon responds to its controversy in human intelligence research but also to the counterintuitive character that it would have for artificial intelligence. If SLODR were true in AI we would have that as long as we construct more general-purpose AI systems, we would have that they show less correlation in perfor-

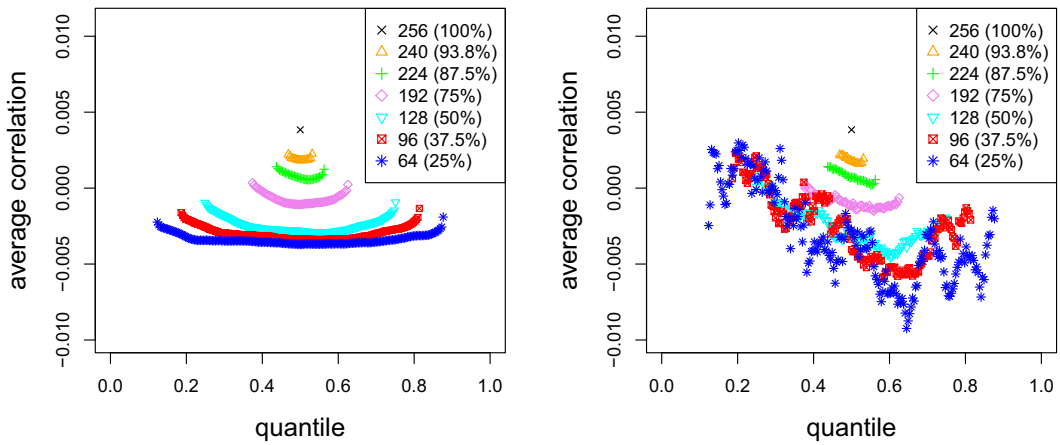


**Figure 5.** Average correlations for 256 agents (randomly-generated programs) and 256 tasks (all the ECA rules), using mirrored rewards for half of the trials). Left: average correlation per quantiles using several bin sizes, where results are sorted by agent performance. Right: each bin is only evaluated with the tasks of that quantile of difficulty.

mance among a range of tasks thus, in a way, becoming less general.

We have not only challenged that possibility but also analysed whether a reverse universal law of augmenting returns (ULOAR) may appear in a very simple setting of tasks and agents, even with agents of very reduced ability. Closely related to the methodological issues of the analysis of SLODR in human intelligence, we have seen how important it is to perform the analysis in the right way, by using difficulty to have an appropriate level of challenging tasks for each subpopulation to account for the variances. The advantages of this artificial experimentation setting is that we can rule out many other confounding factors that appear in human intelligence, such as the existence of some tasks that have been more common in our evolutionary history or culture, the existence of more efficient specialised modules in our brain predisposed for them, etc.

It is too soon to see whether the current questions and the methodology used here can have any effect in the way general-purpose AI agents will be developed and evaluated in the future, including multi-agent architectures, for which the environment and policies can be extended relatively easily [20]. However, there are some areas in AI that can benefit from some of the issues raised in this paper more immediately. For instance, the ULOAR can suggest new ways of empirically analysing AI systems, to devise new benchmarks and competitions and, most especially, to analyse their results. Of course, in these cases, the tasks and agents would be less minimal (more realistic), but would have more issues about how arbitrarily they have been chosen: many benchmarks include many tasks for which researchers have specialised during decades, and the agents would be a biased subpopulation composed of the participants of the competitions. Another issue for real competitions would be the estimation of difficulty, which is necessary to make the analysis properly. We advocate for principled approaches, based on the policy descriptions, as done here, but other approaches such as Item Response Theory could be used [15, 7, 37].

Some competitions in AI would be better suited than others to the concept of generality. For instance, while we can understand the notion of generality for a planning competition [36] (i.e., a general planner would be the one that is good for a wide range of planning problems), it is for general-purpose agents where the notion of a general factor is more intuitive and closer to the original notions in human intelligence. For instance, the reinforcement learning competition [54, 10] or the general video game competition [43, 42] would have similar interpretations of results as those discussed here. Nevertheless, the use of correlation matrices for whatever AI competition may show some general factors appearing. We could also investigate whether they grow stronger or not, as the discipline advances.

Finally, an area that can be particularly suitable for this kind of analysis is machine learning. There are already several 'experiment databases' [51, 52] whose results can be used to analyse correlations, positive manifolds and whether SLODR (or ULOAR) is taking place there. The interpretation of the results will likely be intriguing but the scope and implications may be fascinating.

## ACKNOWLEDGEMENTS

We thank the anonymous reviewers for their comments. This work has been partially supported by the EU (FEDER), by Generalitat Valenciana PROMETEOII/2015/013 and Spanish MINECO grant TIN2015-69175-C4-1-R.

## REFERENCES

[1] D. Andre and S.J. Russell, 'State abstraction for programmable reinforcement learning agents', in *Proceedings of the National Conference on Artificial Intelligence*, pp. 119–125, (2002).
[2] M. G. Bellemare, Y. Naddaf, J. Veness, and M. Bowling, 'The arcade learning environment: An evaluation platform for general agents', *Journal of Artificial Intelligence Research*, **47**, 253–279, (06 2013).
[3] C. Boutilier, R. Reiter, M. Soutchanski, S. Thrun, et al., 'Decision-theoretic, high-level agent programming in the situation calculus', in *Proceedings of the National Conference on Artificial Intelligence*, pp. 355–362, (2000).
[4] S. Bringsjord, 'Psychometric artificial intelligence', *Journal of Experimental & Theoretical Artificial Intelligence*, **23**(3), 271–277, (2011).
[5] S. Bringsjord and B. Schimanski, 'What is artificial intelligence? Psychometric AI as an answer', in *International Joint Conference on Artificial Intelligence*, pp. 887–893, (2003).
[6] A. Cangelosi and M. Schlesinger, *Developmental robotics: From babies to robots*, MIT Press, 2015.
[7] Rafael Jaime De Ayala, *Theory and practice of item response theory*, Guilford Publications, 2009.
[8] I. J. Deary, V. Egan, G. J. Gibson, E. J. Austin, C. R. Brand, and T. Kellaghan, 'Intelligence and the differentiation hypothesis', *Intelligence*, **23**(2), 105–132, (1996).

[9] D. K. Detterman and M. H. Daniel, 'Correlations of mental tests with each other and with cognitive variables are highest for low IQ groups', *Intelligence*, **13**(4), 349–359, (1989).

[10] C. Dimitrakakis, G. Li, and N. Tziortziotis, 'The reinforcement learning competition 2014', *AI Magazine*, **35**(3), 61–65, (2014).

[11] D. L. Dowe and A. R. Hajek, 'A computational extension to the Turing Test', in *Proceedings of the 4th Conference of the Australasian Cognitive Science Society, University of Newcastle, NSW, Australia, also as Technical Report #97/322, Dept Computer Science, Monash University, Australia*, (1997).

[12] D. L. Dowe and J. Hernández-Orallo, 'IQ tests are not for machines, yet', *Intelligence*, **40**(2), 77–81, (2012).

[13] D. L. Dowe, J. Hernández-Orallo, and P. K. Das, 'Compression and intelligence: social environments and communication', in *Artificial General Intelligence*, eds., J. Schmidhuber, K.R. Thórisson, and M. Looks, volume 6830, pp. 204–211. LNAI series, Springer, (2011).

[14] B. Edmonds, 'The social embedding of intelligence', in *Parsing the Turing Test*, eds., Robert Epstein, Gary Roberts, and Grace Beber, 211–235, Springer, (2009).

[15] S. E. Embretson and S. P. Reise, *Item response theory for psychologists*, L. Erlbaum, 2000.

[16] G. J. Fogarty and L. Stankov, 'Challenging the "law of diminishing returns"', *Intelligence*, **21**(2), 157–174, (1995).

[17] M. Genesereth, N. Love, and B. Pell, 'General game playing: Overview of the AAAI competition', *AI Magazine*, **26**(2), 62, (2005).

[18] M. Genesereth and M. Thielscher, 'General game playing', *Synthesis Lectures on Artificial Intelligence and Machine Learning*, **8**(2), 1–229, (2014).

[19] J. Hernández-Orallo, 'Beyond the Turing Test', *J. Logic, Language & Information*, **9**(4), 447–466, (2000).

[20] J. Hernández-Orallo, 'On environment difficulty and discriminating power', *Autonomous Agents and Multi-Agent Systems*, 1–53, (2014).

[21] J. Hernández-Orallo, 'C-tests revisited: Back and forth with complexity', in *Artificial General Intelligence - 8th International Conference, AGI 2015, Berlin, Germany, July 22-25, 2015, Proceedings*, eds., J. Bieger, B. Goertzel, and A. Potapov, pp. 272–282. Springer, (2015).

[22] J. Hernández-Orallo, 'Stochastic tasks: Difficulty and Levin search', in *Artificial General Intelligence - 8th International Conference, AGI 2015, Berlin, Germany, July 22-25, 2015, Proceedings*, eds., J. Bieger, B. Goertzel, and A. Potapov, pp. 90–100. Springer, (2015).

[23] J. Hernández-Orallo, 'Evaluation in artificial intelligence: From task-oriented to ability-oriented measurement', *Artificial Intelligence Reviews*, (2016).

[24] J. Hernández-Orallo, *The Measure of All Minds: Evaluating Natural and Artificial Intelligence*, Cambridge University Press, 2016.

[25] J. Hernández-Orallo and D. L. Dowe, 'Measuring universal intelligence: Towards an anytime intelligence test', *Artificial Intelligence*, **174**(18), 1508 – 1539, (2010).

[26] J. Hernández-Orallo, D. L. Dowe, and M. V. Hernández-Lloreda, 'Universal psychometrics: Measuring cognitive abilities in the machine kingdom', *Cognitive Systems Research*, **27**, 5074, (2014).

[27] J. Hernández-Orallo, F. Martínez-Plumed, U. Schmid, M. Siebers, and D. L. Dowe, 'Computer models solving human intelligence test problems: progress and implications', *Artificial Intelligence*, (2016).

[28] M. Hutter, *Universal Artificial Intelligence: Sequential Decisions based on Algorithmic Probability*, Springer, 2005.

[29] C. Igel and M. Toussaint, 'A no-free-lunch theorem for non-uniform distributions of target functions', *Journal of Mathematical Modelling and Algorithms*, **3**(4), 313–322, (2005).

[30] J. Insa-Cabrera, D. L. Dowe, and J. Hernández-Orallo, 'Evaluating a reinforcement learning algorithm with a general intelligence test', in *Current Topics in Artificial Intelligence. CAEPIA 2011*, eds., J.A. Lozano, J.A. Gamez, and J.A. Moreno. LNAI Series 7023, Springer, (2011).

[31] A. R. Jensen, *The g factor: The science of mental ability*, Westport, Praeger, 1998.

[32] Arthur R Jensen, 'Regularities in Spearman's law of diminishing returns', *Intelligence*, **31**(2), 95–105, (2003).

[33] S. Legg and M. Hutter, 'Universal intelligence: A definition of machine intelligence', *Minds and Machines*, **17**(4), 391–444, (2007).

[34] J. Leike and M. Hutter, 'Bad universal priors and notions of optimality', in *Conference on Learning Theory, CoLT*, (2015).

[35] M. Leonetti and L. Iocchi, 'Improving the performance of complex agent plans through reinforcement learning', in *Proceedings of the 2010 International Conference on Autonomous Agents and Multiagent Systems: volume 1*, pp. 723–730, (2010).

[36] D. Long and M. Fox, 'The 3rd international planning competition: Results and analysis', *J. Artif. Intell. Res. (JAIR)*, **20**, 1–59, (2003).

[37] Fernando Martínez-Plumed, Ricardo B. C. Prudêncio, Adolfo Martínez-Usó, and José Hernández-Orallo, 'Making sense of item response theory in machine learning', in *ECAI*, IOS Press, (2016).

[38] J. McDermott, D. R. White, S. Luke, L. Manzoni, M. Castelli, L. Vanneschi, W. Jaśkowski, K. Krawiec, R. Harper, K. De Jong, and U.-M. O'Reilly, 'Genetic programming needs better benchmarks', in *Proceedings of the 14th international conference on genetic and evolutionary computation conference*, pp. 791–798. ACM, (2012).

[39] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis, 'Human-level control through deep reinforcement learning', *Nature*, **518**(26 February), 529–533, (2015).

[40] A. L. Murray, H. Dixon, and W. Johnson, 'Spearman's law of diminishing returns: A statistical artifact?', *Intelligence*, **41**(5), 439–451, (2013).

[41] L. Orseau, 'Asymptotic non-learnability of universal agents with computable horizon functions', *Theoretical Computer Science*, **473**, 149–156, (2013).

[42] D. Perez, S. Samothrakis, J. Togelius, T. Schaul, S. Lucas, A. Couëtoux, J.l Lee, C.-U. Lim, and T. Thompson, 'The 2014 general video game playing competition', *IEEE Transactions on Computational Intelligence and AI in Games*, (2015).

[43] T. Schaul, 'An extensible description language for video games', *Computational Intelligence and AI in Games, IEEE Transactions on*, **6**(4), 325–331, (2014).

[44] R. J. Solomonoff. A preliminary report on a general theory of inductive inference, 1960. Report V-131, Zator Co., Cambridge, Ma. Feb 4, revision, Nov.

[45] R. J. Solomonoff, 'A formal theory of inductive inference. Part I', *Information and control*, **7**(1), 1–22, (1964).

[46] C. Spearman, 'General Intelligence, Objectively Determined and Measured', *The American Journal of Psychology*, **15**(2), 201–92, (1904).

[47] C. Spearman, *The abilities of man: Their nature and measurement*, Macmillan, New York, 1927.

[48] R. J. Sternberg, *Handbook of intelligence*, Cambridge University Press, 2000.

[49] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*, MIT press, 1998.

[50] E. M. Tucker-Drob, 'Differentiation of cognitive abilities across the life span.', *Developmental psychology*, **45**(4), 1097, (2009).

[51] J. Vanschoren, H. Blockeel, B. Pfahringer, and G. Holmes, 'Experiment databases', *Machine Learning*, **87**(2), 127–158, (2012).

[52] J. Vanschoren, J. N. van Rijn, B. Bischl, and L. Torgo, 'Openml: networked science in machine learning', *ACM SIGKDD Explorations Newsletter*, **15**(2), 49–60, (2014).

[53] D. R. White, J. McDermott, M. Castelli, L. Manzoni, B. W. Goldman, G. Kronberger, W. Jaśkowski, U.-M. O'Reilly, and S. Luke, 'Better GP benchmarks: Community survey results and proposals', *Genetic Programming and Evolvable Machines*, **14**, 3–29, (2013).

[54] S. Whiteson, B. Tanner, and A. White, 'The Reinforcement Learning Competitions', *The AI magazine*, **31**(2), 81–94, (2010).

[55] S. Wolfram, *A new kind of science*, Wolfram media, Champaign, IL, 2002.

[56] D. H. Wolpert, 'What the no free lunch theorems really mean; how to improve search algorithms', Technical report, Santa fe Institute Working Paper, (2012).

[57] D. H. Wolpert and W. G. Macready, 'No free lunch theorems for search', Technical report, SFI-TR-95-02-010 (Santa Fe Institute), (1995).