Analogy as the Core of Intelligence

Akira Pyinya

akirapyinya@gmail.com

Abstract

This article argues that the core of intelligence is not optimization, but analogy. We define intelligence as "doing the same thing as the examples of the right thing to do in new situations." We transform Hofstadter's Copycat problem into a sequence prediction problem to derive a formal definition of analogy-based intelligence, from which value functions and temporal-difference error can be derived, showing that optimizers can be derived from analogy-based systems. We demonstrate how agency and free will arises from conflicts between different predictions based on different examples.

1. Optimization and Analogy

There are two different approaches to intelligence: optimization and analogy. To compare them, we can start with the following argument: *An intelligent agent is a system that does the right thing* (Russell 1991) .

In order to do the right thing, an intelligent agent must first know what the right thing is. For example, the right thing for organisms to do is to survive and reproduce, while the right thing for a robot vacuum to do is to clean the room. However, this is only a general idea, not a practical description. The universe is vast, complex, and sometimes unpredictable, it is impossible to list every right thing to do in the universe. Therefore, we must simplify this list to fit the agent's limited memory space.

This list can be simplified in two ways: *functions* and *examples*.

From the viewpoint of functions, a general metric can be applied to the entire world. With this metric, the rightness of things can be represented by functions such as reward functions or value functions. For example, to do the right thing, a robot vacuum should maximize the reward obtained from cleaning, or optimize a cleanness function. Optimization is the engine of function-based intelligence.

On the other hand, the right thing to do can be represented by examples. For example, a neural network recognizes digits in the MNIST dataset by learning from training datasets, which consists of examples of mapping groups of pixels to labels. The agent tries to do the same thing like the examples, in other words, make analogies. Analogy is the engine of example-driven intelligence.

We can compare optimization approaches with analogy approaches.

1.1. Optimization: Advantage and Limitation

From the perspective of optimization, every right thing has a degree of rightness that can be measured, ranked, and improved. If there is a right thing to do, it's possible to do it better.

An optimizer always tries to do better. For example, if making one paperclip is good, then making two paperclips or making a paperclip in a shorter time and/or cost less is better. Increasing total human happiness is better than making one person slightly happier.

An ideal optimizer should behave like *Homo economicus*, trying to maximize expected rewards, utility, or some other measurements. Utility functions or value functions estimate the degree of rightness, so maximization is the way leads to the most possible right thing to do.

1.1.1 Optimizers are flexible

One advantage of optimizers is their flexibility. Functions can be universal, in other words, they can be applied to every possible situation in the world, or at least every situation in which the agent is supposed to operate, including those the agent has never encountered before.

Flexibility is important because a system focused on one specific task cannot be very intelligent. Artificial General Intelligence (AGI) is supposed to be capable of performing any type of task or pursuing any final goal. According to the orthogonality thesis, any level of capability can be combined with any final goals (Bostrom 2012). Intelligence is believed to be the capability of optimization, the measurable ability of doing better in all circumstances (Legg & Hutter 2007).

1.1.2 Optimization has a formal definition

The optimization problem can be defined as finding the optimal solution within the search space according to some objective function (Valencia-Rivera et al., 2024). Optimization-based intelligence is formally defined by AIXI (Hutter 2003) . AIXI takes the best possible action to maximize the expected total rewards. AIXI is based on Solomonoff induction, which is the theoretically best way to predict a sequence, though it is not computable. Nevertheless, AIXI's computable approximation is arguably equivalent to optimization-based intelligence in the real world.

1.1.3 But real-world intelligence may not be optimizers

From the viewpoint of optimization, an agent has a core hidden deep within that is like "an unchanging indivisible atom" (Demski & Garrabrant 2019). In this core we can write down a "final goal" such as maximizing total rewards (Omohundro 2018). Once the core is sealed and the start button is pressed, the agent will always follow the command from the core. This model is simple and intuitive, but may not be a good model of real-world intelligence.

In theory, specifying a final goal is as simple as typing a command on a computer. However, in the real world, it is difficult to translate a goal into code correctly (Bostrom 2012). Even worse, we have no idea about what our own "utility function" is, except that it is extremely complex, and we must align machines' utility function with ours, "any mistake in objective may have negative consequences." (Russell 2016)

Some people even claim that assigning any goal to a "superintelligence" would lead to disastrous consequences, because the superintelligence would pursue that goal with all the available resources, such as transforming our solar system into a "computronium" (Bostrom 2012). Therefore, the problem is not the wrong goal, but the optimizer model itself.

In conclusion, real-world intelligence agents, such as humans, do not have fixed goals, and cannot be properly described by the final goal hypothesis. Optimizers may not be the best model for real world intelligence.

But, do we have other options?

1.2. Analogy: Normativity by Examples

"Doing the right thing" can also be guided by examples, because information about the right thing to do can be found in examples of the right thing to do. Through analogy, old and local examples can serve as a universal guide for new environments without the need for distillation into functions.

1.2.1 Analogy as "indirect normativity"

Many reinforcement learning approaches use analogies to speed up the learning processes. However, analogies only play a heuristic role, as aid to discovery. These approaches fall under the umbrella category of "indirect normativity"(Bostrom 2012). Examples of indirect normativity include *imitation learning* and *value learning*.

Imitation learning mimics the behavior of examples to improve performance. It's efficient, relatively safe, and can learn to do things that are hard to describe by goals, but struggles in environments it has never encountered (Christian 2021).

Value learning approaches, such as Inverse Reinforcement Learning (IRL), attempt to extract a value function from behavioral examples (Ng & Russell 2000). For example, when given a set of human-generated driving data, IRL tries to extract an approximation of humans' value function of driving. IRL approaches are more flexible to be applied in novel situations, but still need to assign a value function to the agent, which hits the same brick wall as the direct specification approaches.

However, there may be ways to bypass the value function. If we can estimate the value function, in other words, how good it is for the agent in a given state under a given policy, why don't we calculate the best policy for the current state by request, without reconstructing the value function for all the states?

Autonomous driving vehicles, for example, face an endless list of "long-tail" cases, such as earthquakes or the vehicle catching fire. It would take too many computational resources to estimate all of their rewards beforehand. Can we make real-time estimations of the right thing to do in new "long-tail" cases based directly on examples, without referring to value functions?

1.2.2 Following the examples directly

A good example of "doing the right thing by following examples" is the common law system. Common law courts make decisions by "following the precedent" rather than relying primarily on the written code of law (Postema 2004), because "if there are good reasons to believe that an earlier case was correctly decided, and if the facts in a later case are the same as those in the earlier case, then there are good reasons for believing that the same decision would be correct in the later case. "(Lamond 2016)

Artificial intelligence follows examples in a similar way. Machine learning, for example, is known as data-driven. In other words, actions are taken based on examples rather than fixed functions. Information about "the right thing to do" is learned from examples in training datasets. For example, Large Language Models (LLM) respond to human questions in the same way as the language material they have learned.

Performance measures are also hidden within datasets. A learning program should run on a test dataset to evaluate its performance. In principle, the program should be unaware of the test dataset during the training process, that's why it is impossible to distill a performance measure function for training.

Natural intelligence is also driven by examples. Natural selection preserves the memories of doing the right thing in the gene pool, and removes those that are not. According to Assembly Theory, life follows "a lineage of events stemming from the origin of life"(Walker 2024), which contains a memory of billions of years' of evolutionary history, during which all the events were "the right thing to do". If one of these events went wrong, in other words, if it led to a failure to survive and reproduce, the organism wouldn't exist today. Organisms do the right thing by following examples, "If your past has more possibilities, your future has more" (Walker 2024) . In other words, *intelligence is the self-augmenting memories of doing the right thing*. The intelligent agent is its memory.

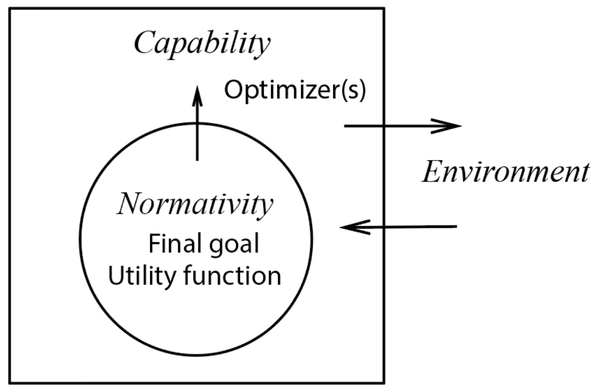

FIGURE 1. In traditional intelligence models, the normativity is a final goal that can be assigned to an agent, and can be combined with optimizers of different capabilities.

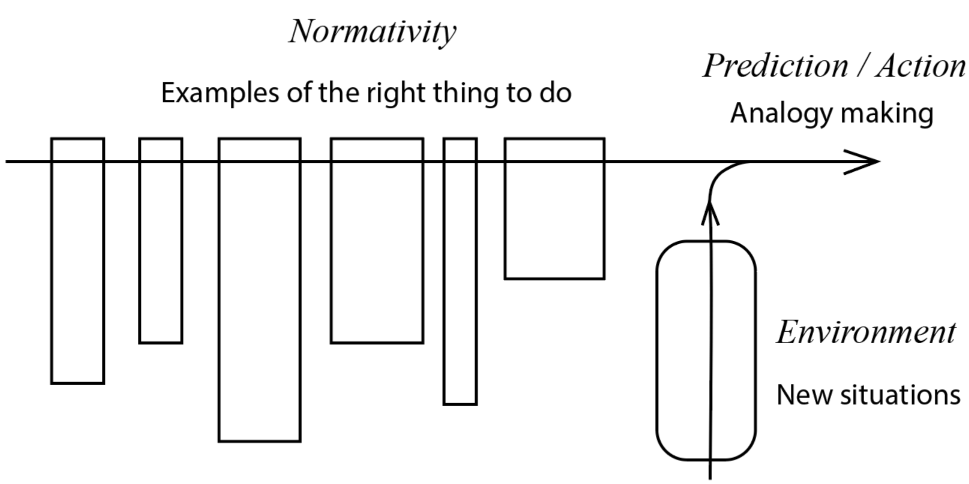

FIGURE 2. In our model, actions in new situations are driven by predictions based on examples of the right thing to do.

1.2.3 Why analogy

Analogy-based intelligence is free from the Paperclips Maximizer Paradox. Optimizers may pursue a fixed goal at all costs, which can result in catastrophe, whereas analogy-based intelligence would not pursue a fixed goal at all costs if "doing things at all costs" is not included among the examples of the right thing to do, because doing one thing at all costs is radically different from the examples. For example, using all the world's resources to make paperclips is not the right thing to do, because "using all the world's resources" is not similar to the actions in the examples, unless you deliberately train the system with examples of depleting resources and destroying environments. Furthermore, analogy-based intelligence can change its goal by adding or deleting examples from its library of examples, whereas optimizers rely on fixed goals.

More importantly, we argue that analogy is a better model of intelligence, and optimization is merely a conventional tool. Analogy is not only the core of cognition (Hofstadter 2001), but also the core of intelligence.

2. Formalize Analogy-based Intelligence

In summary, analogy-based intelligence can be defined as:

Definition 1*: Intelligence is doing the same thing in new situations as examples of doing the right thing.*

But what does it mean "doing the same thing in different situations"? We have no clear definition yet.

In contrast, the optimization camp has a formal definition of intelligence called AIXI. AIXI defines the optimization problem by converting it into a sequence prediction problem. According to Solomonoff induction, sequence prediction problems have a universal solution: every letter string is considered the output of a Turing machine, and because the prior probability of all possible Turing machines can be estimated, the next letter of any letter string can also be estimated (Hutter 2003, Legg 1997, Solomonoff 1964).

We can define analogy-based intelligence in a similar way. If we can rewrite "how to do the same thing as the examples" as a sequence prediction problem, we can rephrase it in the language of Solomonoff induction, and construct a formal definition of analogy-based intelligence.

2.1 Copycat in the language of sequence prediction

We begin with the Copycat project, an analogy-based intelligence approach that formalizes analogy by asking a question:

Question 1: *Suppose the letter-string* ***abc*** *were changed to* ***abd****; how would you change the letter-string* ***yk*** *in the same way*? (Hofstadter & Mitchell 1995)

We can ask a question similar to Question 1:

Question 2 : *Suppose there is a very long letter-string, in which every fragment* ***abc*** *is followed by an* ***abd****. At the end of the string there are letters* ***yk****. What will be the next letters?*

Question 2 is a sequence prediction problem that aims to estimate the continuation of a sequence.

....
dsiouyu**abcabd**kvlhpwy
gjzbmfjqupghckahwmvb
abcabdkslnvofhdk
qucjslcmzlapmfhsnclwpr
ssfcddffsjhkncismsdsfew
edvdsoopr**abcabd**purcs
asdwqsjjfkrk**yk**????
....

FIGURE 3. The Copycat question is transformed into a sequence prediction problem: How to predict the fragment "????" at the end of a very long letter-string?

2.2 Formalize analogy using Solomonoff induction

In Question 2, the example of the right thing to do, "*every fragment* ***abc*** *is followed by an* ***abd****"* , describes a probability distribution over all possible sequences. We can ask more formally:

Question 3: *A set of examples of doing the right thing is represented by a probability distribution* p_{ER} *over all possible history* $h \in H$ *at a given moment. Let* $f \in F$ *represents a possible future, what is f's probability of being the right thing to do?*

A sequence's probability of being the right thing to do can be estimated using Solomonoff induction: if *h* represents the history of doing the right thing, the continuation of *h* is the future most likely doing the right thing, too, because the right thing to do in the future is most likely to be the continuation of the right thing to do in the past. Thus the analogy question is transformed into a sequence prediction question.

Solomonoff induction can estimate the prior probability of any sequence, known as Solomonoff prior, which is inverse proportion to the Kolmogorov complexity, the length of the shortest Turing machine that can generate this sequence. Solomonoff prior is uncomputable. The probability of a future f being the continuation of a history *h* is in proportion to the Solomonoff prior of their combination *hf* divides the Solomonoff prior of *h* (Legg 2004):

Theorem 1: *Given h is a sequence represents the history in which an intelligent agent is doing the right thing, f is a sequence that represents a future , f's probability of being the right thing to do is:*

$$p_R(f) \propto M(hf)/M(h)$$
$$p_R(f) \propto K(h)/K(hf)$$

M stands for Solomonoff prior and K stands for Kolmogorov complexity.

It's a formal way to say that intelligence is both a prediction problem (Clark 2024, Friston et al., 2010) and a data compression problem. Using Theorem 1, we can further estimate the possible future based on a distribution of the right thing to do:

Theorem 2: *A set of examples of doing the right thing is represented by a probability distribution* p_{ER} *over all possible history* $h \in H$ *at a given moment. Let* $f \in F$ *represents a possible future, f's probability of being the right thing to do is the weighted sum of its probability of being the concatenation of all h:*

$$p_R(f) \propto \sum_{h \in H} p_{ER}(h) M(hf)/M(h)$$

$$p_R(f) \propto \sum_{h \in H} p_{ER}(h) K(h) / K(hf)$$

2.3 Derive value functions from analogy

We argue that analogy is not just an aid of optimization. Analogy is the core of intelligence. Like the theory of relativity is more general than Newtonian mechanics, the analogy-based model is more general than the optimizer model. Using Theorem 2, we can derive the utility functions, the basis of the optimizer model, from the analogy-based model.

2.3.1 Value of a possible future

Reinforcement learning uses value functions to estimate how good it is for the agent according to some variable such as a state under a policy (Sutton & Barto 2018). This value is often measured as the sum of discounted future rewards.

From the perspective of analogy, the world is not divided into "the agent" and "environment", so we neither divide the world into "states" and "policies", nor do we draw an arbitrary boundary between "standard input" and "reward input" as AIXI does. (We will revisit the free choice problem in chapter 3).

Analogy-based intelligence directly estimates how good a future is, in other words, *a possible future's probability of being the right thing to do*. If we define value as the estimation of how good something is, this probability is *the value of a possible future*.

We can estimate the value of a possible future by rewrite Theorem 1 as:

Theorem 3: *Given h is a sequence represents the history in which an intelligent agent is doing the right thing, f is a sequence that represents a future, the value of f is :*

$$v_R(f) \propto M(hf) / M(h)$$
$$v_R(f) \propto K(h) / K(hf)$$

And from Theorem 2 we can have value functions based on a distribution of the right thing to do:

Theorem 4: *A set of examples of doing the right thing is represented by a probability distribution* p_{ER} *over all possible history* $h \in H$ *at a given moment. Let* $f \in F$ *represents a possible future, f's value is the weighted sum of its probability of being the concatenation of all h:*

$$v(f) \propto \sum_{h \in H} p_{ER}(h) M(hf)/M(h)$$

$$v(f) \propto \sum_{h \in H} p_{ER}(h) K(h)/K(hf)$$

2.3.2 Reward is not needed

Traditional reinforcement learning is based on *the reward hypothesis*, "all of what we mean by goals and purposes can be well thought of as maximization of the expected value of the cumulative sum of a received scalar signal (reward)" (Sutton & Barto 2018) An agent is assumed to receive a reward signal that is distinct from normal observations. The discounted sum of rewards is assumed to be the value to maximize.

However, the reward signal has no physical basis; rather, it is "an abstraction summarizing the overall effect of a multitude of neural signals generated by many systems in the brain that assess the rewarding or punishing qualities of sensations and states" (Sutton & Barto 2018). Furthermore, the reinforcement signal does not have to be the reward signal.

Analogy-based models can estimate the value function directly through analogy, and find it difficult and unnecessary to attribute value to a reward signal that is observed every moment. This is why there's no place for the reward signal in the analogy-based model.

2.3.3 Value functions of distributions over possible futures

Either states of the world or a state under a policy can be described as a probability distribution over all possible futures ("now" is included in the futures). So using Theorem 3, we can also estimate the value function of a state, or that of a state under a policy, or any other conditions that can be described by a probability distribution over all possible futures:

Theorem 5: *A set of examples of doing the right thing is represented by a probability distribution p_{ER} over all possible history $h \in H$ at a given moment. A condition C is described by p_C, a probability distribution over all possible futures. The value function of condition C is the weighted sum of all possible futures' value:*

$$v(C) \propto \sum_{h \in H, f \in F} p_{ER}(h) p_C(f) M(hf)/M(h)$$

$$v(C) \propto \sum_{h \in H, f \in F} p_{ER}(h) p_C(f) K(h)/K(hf)$$

We have defined a value function of a state under a policy, should we go on optimizing the policy in the same way as the traditional reinforcement learning approach?

2.3.4 New examples and ever changing value functions

However, it is problematic to optimize this value function directly. Unlike traditional value functions, analogy-based value functions can change over time, because agents are constantly creating new examples of doing the right thing.

As time passes, possible futures solidify into pasts, and events that are remembered become new examples of the right thing to do. The agent collects these new examples and updates its prediction of the right thing to do in the future based on this updated library of examples. According to Theorem 3, the agent constantly updates its value function. The same possible future may have different values at different moments.

Another issue with policy optimization is the lack of a definition of "choosing." In Theorem 3, the possible futures are considered as a whole, without distinguishing between the agent and the environment. Thus, there is no distinction between states and policies, nor presumption of agency. How should we describe choosing between two policies? Is this choice also a policy? In an analogy-based model, we must reveal the meaning of agency.

3. The Emergence of Agency

Agency is the default assumption for many intelligence models. For example, AIXI assumes that an agent can choose from different options to maximize total rewards, because agency is the basis of optimization by definition.

In contrast, analogy-based models start with a world without agency. There is no preset boundary between an agent and the environment, so we cannot even distinguish the input from the output. The future is predicted based on the past without stating that the meaning of the prediction is "the agent should do" or "the agent will do".

However, agency is necessary for understanding intelligence. An intelligence model should explain agency, if not begin with it. We feel that we are making choices, we feel that agents are pursuing goals. We must demonstrate how agency and mind-body dualism emerge from a model that begins without them.

3.1 Learn how to survive from survivorship bias

One problem with data-driven induction is the survivorship bias. The observer has no access to experiences involving its own death, such as "an agent dropping an anvil on its own head". Based on experiences, an agent cannot predict what would happen if an anvil were dropped on its head, because it has no experience with a destroyed head (Demski & Garrabrant 2019). Hohwy (2016) argues that "we cannot obtain an independent view of our position in the world." The viewpoint of the agent is imprisoned in its own body.

Although survivorship bias can hinder our understanding of the world, it can also teach us how to survive—the right thing to do for organisms. One can learn how to survive by studying survivorship bias. For example, we can confidently predict that an organism is highly unlikely to drop an anvil on its own head, because organisms' actions can be predicted based on what they and their lineage did when they were alive. We know that neither they nor their lineage ever did so. Therefore, experiences that are excluded by survivorship bias are not the right thing to do.

3.2 Understand the world through theory of mind

The question remains. Advanced intelligent agents should be able to step outside their own perspective and develop "the view from nowhere", which is an objective view of themselves (Nagel 1986). How can you correctly predict the entire world based only on examples of the right thing to do?

The theory of mind provides an indirect way to overcome the survivorship bias. Advanced intelligent agents can speculate about others' perceptions without perceiving them directly (Frith & Frith 2005).

Through experience, agents can develop the theory of mind by understanding that other agents are similar to themselves. The same models can be used to predict both itself and other agents. Thus, an agent can speculate about its own death from others' viewpoints by observing the death of others. The agent can then conclude that similar things will happen if it drops an anvil on its own head. .

In other words, with the theory of mind, agents can learn that the same laws of physics apply to their own body as to the bodies of others, and understand their own body from an objective view.

3.3 Agency arise from the conflict between two predictions

3.3.1 The subjective prediction and the objective prediction

Now we have two different ways of making predictions: the internal way and the external way. The internal way makes predictions based on the examples of the right thing to do, which contain survivorship bias due to selection. The external way fixes this bias by using the indirect knowledge obtained through the theory of mind. We can call the internal way *the subjective prediction* and the external way *the objective prediction*.

There are conflicts between these two predictions, and most of them occur in and around an agent's own body. The subjective prediction suggests that the agent is highly unlikely to die. It predicts that an agent cannot drop an anvil on its own head because it has no experience of such an event. On the contrary, the objective prediction suggests that the same laws of physics apply to one's own body as to the external world. Therefore, it may predict some risk of dropping an anvil on one's own head, because doing so is not forbidden by the laws of physics. In other words, the subjective prediction describes a world in which the agent always survives, while the objective prediction describes a world that is independent of the agent's survival.

3.3.2 Wantings are conflicts between predictions

What are conflicts between predictions like? Imagine two tennis players, A and B, playing against each other. Player A is trying to win a point, in other words, she predicts that she will win the point based on her actions, perceptual information, and memories. And vice versa for Player B. However, the two predictions cannot both be correct. The player with more useful information, more computational resources, and better prediction techniques will make better predictions and win more points. In this case, we can say that both the two players *want to win*. The player better at understanding others' predictions and using it to improve her own prediction is more likely to win.

We can define *wanting* as "the conflict between predictions from different perspectives." We feel that an agent wants to change the world to achieve its goal, because there are conflicts between its subjective prediction and objective prediction.

For example, one day, you are driving to work and everything seems fine. Suddenly, a car approaches from the opposite direction and about to collide. Based on your experience watching traffic accidents and other physical movements, you can make an objective prediction that there is a high risk of collision and a threat to life. On the contrary, a subjective prediction would indicate that the right thing to do is to avoid the collision and survive. Then, your body and brain behave as if you want to avoid the collision. If the subjective prediction is proved correct, you will feel like you have overcome a challenge and saved your life. Otherwise, you will feel as if you failed to prevent a disaster, if you can still think and feel at all.

3.3.3 The right thing to do is to resolve conflicts

In summary, when the subjective prediction is proved correct, we feel like we have overcome the obstacles and achieved our goals. Otherwise, we will feel defeated by the world. Furthermore, we can improve our model to align the two predictions with each other. For example, people feel as if natural forces are against them until they learn techniques to make use of the forces to achieve their goals. A driver who spends several days on chaotic roads may learn to avoid collisions more easily and make her objective predictions align better with subjective predictions.

This explains the Active Inference's argument that an agent minimizes its surprise by either changing the world or changing its mind (Friston et al., 2017). From the perspective of prediction conflicts, the surprise that the agent minimizes is the conflict between the objective prediction and the subjective prediction. Resolving these conflicts is the right thing to do for an agent.

When there are a lot of conflicts between predictions, we feel like that we want so badly, and sometimes we feel as if the environment is working against us. If there are fewer or no conflicts, we feel serene and want less. A significant conflict between a potential collision and survival represents an emergency for drivers. Experienced drivers and/or driving in a more peaceful traffic will experience fewer conflicts and less panic.

However, there is no clear or fixed boundary between the internal self and the external environment. The boundary of the agent is not its "skin", but rather an "extended mind."(Clark & Chalmers 1998) Figuring out this boundary is similar to the Serenity Prayer about "grace to accept with serenity the things that cannot be changed, courage to change the things which should be changed, and the wisdom to distinguish one from the other." It's the wisdom of distinguishing subjective predictions from objective predictions.

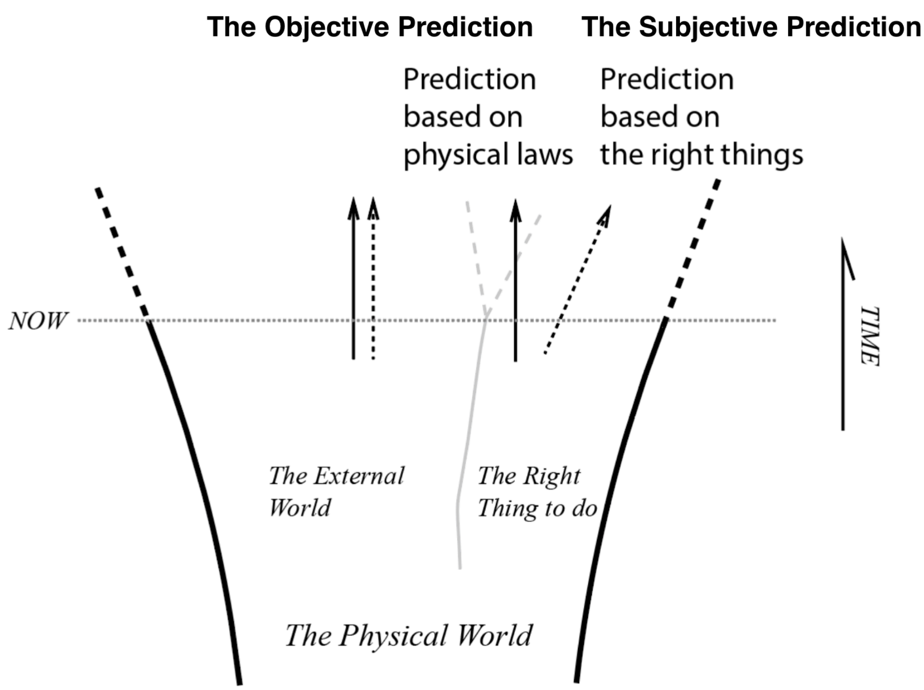

FIGURE 4. An agent would make two different predictions about the world: the objective prediction is based on the laws of physics, while the subjective prediction is based on the examples of the right thing to do. The conflict between these two predictions is the gap between the external world and our wantings.

3.4 Free will and conflicts between predictions

However, in most cases of conflicts, different predictions do not yield opposite results. Instead, subjective predictions provide details that reduce the uncertainties in objective predictions.

For example, I just made a cup of tea. Neither making tea nor making coffee violates the laws of physics, in other words, the objective prediction does not specify which drink I would make. However, the subjective prediction indicates that making tea aligns with my memories of the right thing to do, from my childhood experience with different drinks to everything I have done so far today. The subjective prediction provides additional detailed information that the objective prediction cannot.

According to Seth (2021), having free will means that an action should come from within, and that one could have done otherwise. In the case of making tea, objective predictions do not specify which drink to choose, which means I could have chosen otherwise. Subjective predictions, on the other hand, come from within and provide more details about my choice of drink. Thus we can conclude that the conflict between subjective prediction and objective predictions is the origin of free will.

3.4.1 Fewer examples, more detailed predictions

Some may wonder how a subjective prediction, based on fewer examples, could provide more detailed information than an objective prediction based on more examples. In fact, predictions based on selected examples can provide more information than those based on unselected ones.

Imagine there are 100 black balls and 100 white balls in a big box. Inside the box, there is a small bag containing 9 black balls and 1 white ball, which are included in the total of 100 black balls and 100 white balls. To predict the color of a ball randomly chosen from the small bag, we can use 1. fewer examples of the colors of the balls in the small bag (90% black, 10% white) or 2. more examples of the color of the balls in the big box (50% black, 50% white). Option 1 provides more information and makes better predictions.

4. Build an Analogy-based Intelligence

Our model argues that all intelligence agents are analogy-based in theory. However, with the help of analogy-based models, we can build more general and more efficient intelligence, with reinterpreted reinforcement learning tools. Reinforcement signals can be developed without a hypothetical reward signal. Instead, value functions and agency can be derived from analogy.

4.1 Value of the objective prediction

Using the value of objective predictions, we can estimate how good a future is. Using Theorem 5, the value of the future given by an objective prediction P_O is (**Theorem 6**):

$$v(P_O) \propto \sum_{h\in H, f\in F} p_{ER}(h) p_O(f) M(hf)/M(h)$$

We call $v(P_O)$ *the value of an objective prediction (VOP)*. It specifies how good everything is for the agent in the long run, with all available knowledge including those directly observed and those obtained through the theory of mind.

Due to computational limitations, real-world intelligent agents cannot make the best estimation of VOP. In this article, however, we assume that they can, in order to demonstrate the upper bound of their performance.

VOP is always changing, because the agent is constantly updating its objective prediction AND the library of examples of the right thing to do. An increasing VOP indicates that things are getting better, and a decreasing VOP indicates that things are getting worse.

4.2 Analogy-based Temporal-Difference(TD) error

The changing of VOP can be used as a reinforcement signal, allowing us to reinterpret and simplify traditional reinforcement learning methods, such as the temporal-difference (TD) method and actor-critic learning. Additionally, VOP explains how value functions change over time, a topic that traditional models do not address.

The traditional TD method is guided by TD error, which signals discrepancies between current and earlier expectations of reward in the long-term (Sutton & Barto 2018). Although the changing of VOP does not involve rewards, it also signals discrepancies between current and earlier expectations of doing the right thing in the long-term, therefore, it could be a substitute for TD error. From the perspective of analogy, agents appear to be driven by positive TD error signals, because the objective prediction is getting closer to the subjective prediction when an agent is doing the right thing.

According to Theorem 6 we can define TD error based on VOP:

$$\delta = v(P_O) - v(P_O') \propto \sum_{h\in H, f\in F} p_{ER}(h)p_O(f)M(hf)/M(h) - \sum_{h\in H, f\in F} p_{ER}'(h)p_O'(f)M(hf)/M(h)$$

Consider the reward prediction error hypothesis of dopamine neuron activity, which proposes that "one of the functions of the phasic activity of dopamine-producing neurons in mammals is to deliver an error between an old and a new estimate of expected future reward to target areas throughout the brain"(Sutton & Barto 2018). This hypothesis explains why dopamine neurons respond to unpredicted rewarding events. From the perspective of VOP, unpredicted rewarding events indicate a change in the agent's objective prediction and its value, the VOP. Dopamine neurons respond to "neutral cues that precede a reward" because these cues change the agent's objective prediction and VOP as soon as the agent learns their relationship with "rewarding events".

"Rewarding events" can be reinterpreted as events that are more likely to be the right thing to do. For example, getting some food is considered as a rewarding event for animals. However, from the perspective of analogy, it's not a single event, but rather a series of events including seeing or smelling the food, consuming food, and digesting food, etc. Most of these events are likely to be the right thing to do for an animal. The smell of food or a learned neutral cue could predict these events, indicating an increased probability of doing the right thing, in other words, an increased VOP.

Unlike the traditional TD method, VOP reveals another cause of TD error: new examples of the right thing to do. The same predicted future may have different value for the same agent when new examples are introduced. For example, the future of having the same amount of food would have a different value for an animal when it's hungry versus when it's well-fed, because their different histories contain different examples of the right thing to do. Traditional reinforcement learning can explain this difference by introducing a hunger parameter, which is a function of food intake history and some other factors. This function may be more complicated for complex animals. However, analogy-based models offer a universal solution. They interpret all types of value change by directly estimating value based on the historic right thing to do, without introducing additional status parameters.

With VOP derived TD error signals, we can construct agents that can learn to do the right thing using traditional TD methods such as the Actor-Critic method.

5. Conclusion

We demonstrate that analogy-based approaches can be defined more elegantly than the optimization-driven models. Through the lens of analogy, we integrate Algorithmic Information Theory, the Copycat approach, Active Inference, the theory of natural selection, Assembly Theory, and the theory of common law into a big picture of intelligence. Analogy-based approaches avoid paradoxes such as the Paperclip Maximizer problem, and may help us develop safer and more powerful AI.

More importantly, analogy provides a new way of thinking about human intelligence. Humans are not purely rule-based or goal-oriented optimizers. Unlike machines that act automatically, humans behave autonomously (Bates 2024). Our behavior is subject to ourselves, but what is "ourselves"?

From the perspective of analogy, we are self-augmenting memories of our own experiences and those of our evolutionary lineage . Our actions are driven by predictions based on these memories, and are recorded as new memories. We explore the larger world by reinterpreting old memories and creating new ones through increasingly complex interactions with the world. With more and more powerful memory-thinking units, we blaze new trails into the unknown, instead of falling into singularity.

References

Bates, D. W. (2024). An Artificial History of Natural Intelligence: Thinking with Machines from Descartes to the Digital Age. In *An Artificial History of Natural Intelligence*. University of Chicago Press.

Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford University Press.

Christian, B. (2021). *The alignment problem: How can machines learn human values?*. Atlantic Books.

Clark, A., & Chalmers, D. (1998). The extended mind. *analysis*, *58*(1), 7-19.

Clark, A. (2024). *The experience machine: How our minds predict and shape reality*. Random House.

Demski, A., & Garrabrant, S. (2019). Embedded agency. *arXiv preprint arXiv:1902.09469*.

Friston, K. J., Daunizeau, J., Kilner, J., & Kiebel, S. J. (2010). Action and behavior: a free-energy formulation. *Biological cybernetics*, *102*, 227-260.

Friston, K. J., FitzGerald, T., Rigoli, F., Schwartenbeck, P., & Pezzulo, G. (2017). Active inference: a process theory. *Neural computation*, *29*(1), 1-49.

Frith, C., & Frith, U. (2005). Theory of mind. *Current biology*, *15*(17), R644-R645.

Hofstadter, D. R., & Mitchell, M. (1995). The copycat project: A model of mental fluidity and analogy-making. *Advances in connectionist and neural computation theory*, *2*, 205-267.

Hofstadter, D. R. (2001). Analogy as the core of cognition. *The analogical mind: Perspectives from cognitive science*, 499-538.

Hohwy, J. (2016). The self‐evidencing brain. *Noûs*, *50*(2), 259-285.

Hutter, M. (2003). A gentle introduction to the universal algorithmic agent AIXI. *Artificial General Intelligence*.

Lamond, G.(2016). Precedent and Analogy in Legal Reasoning, *The Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/archives/spr2016/entries/legal-reas-prec/.

Legg, S. (2004). *Solomonoff induction*. Department of Computer Science, The University of Auckland, New Zealand.

Legg, S., & Hutter, M. (2007). A collection of definitions of intelligence. *Frontiers in Artificial Intelligence and applications*, *157*, 17.

Nagel, T. (1989). *The view from nowhere*. Oxford university press.

Ng, A. Y., & Russell, S. (2000). Algorithms for inverse reinforcement learning. In *Icml* (Vol. 1, No. 2, p. 2).

Omohundro, S. M. (2018). The basic AI drives. In *Artificial intelligence safety and security* (pp. 47-55). Chapman and Hall/CRC.

Postema, Gerald. (2004). Philosophy of the Common Law. in *Oxford Handbook of Jurisprudence and Philosophy of Law* (pp.588-622)Oxford University Press.

Russell, S. J., & Norvig, P. (2016). *Artificial intelligence: a modern approach*. Pearson.

Russell, S. J., & Wefald, E. (1991). *Do the right thing: studies in limited rationality*. MIT press.

Seth, A. (2021). *Being you: A new science of consciousness*. Penguin.

Solomonoff, R. J. (1964). A formal theory of inductive inference. Part I. *Information and control*, *7*(1), 1-22.

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. Cambridge: MIT press.

Valencia-Rivera, G. H., Benavides-Robles, M. T., Morales, A. V., Amaya, I., Cruz-Duarte, J. M., Ortiz-Bayliss, J. C., & Avina-Cervantes, J. G. (2024). A systematic review of metaheuristic algorithms in electric power systems optimization. *Applied Soft Computing*, *150*, 111047.