Learning Agents with Evolving Hypothesis Classes

Peter Sunehag and Marcus Hutter
{Peter.Sunehag, Marcus.Hutter}@anu.edu.au

Research School of Computer Science
Australian National University
Canberra Australia

Abstract. It has recently been shown that a Bayesian agent with a universal hypothesis class resolves most induction problems discussed in the philosophy of science. These ideal agents are, however, neither practical nor a good model for how real science works. We here introduce a framework for learning based on implicit beliefs over all possible hypotheses and limited sets of explicit theories sampled from an implicit distribution represented only by the process by which it generates new hypotheses. We address the questions of how to act based on a limited set of theories as well as what an ideal sampling process should be like. Finally, we discuss topics in philosophy of science and cognitive science from the perspective of this framework.

1 Introduction

Bayesian inference [BT92,Rob94,Jay03] is arguably the most prominent formal mathematical framework for learning. It is being used in Artificial Intelligence (AI) [RN10] to create intelligent agents, in economics [Sav54,Nya93,OR94] as a model of rational agents, in cognitive science [Edw68,And91,GKT08,TKGG11] as a model of human learning and in philosophy of science [Goo84,HU05,Str06] as a model for induction and scientific progress. The basic principle is that we have an a priori distribution over a hypothesis space and after gaining experience the prior weights are updated using Bayes rule to a posteriori beliefs. Decisions are then taken based on maximum expected utility with respect to the posterior distribution. If an agent performs this procedure sequentially, its a Bayesian reinforcement learning agent [Hut05]. In some of the contexts mentioned above, e.g. philosophy of science, some of the aspects like either the decision making or the sequential nature is often suppressed, while still important for the relevant problem setting.

These are ideal models and can often, as models, fail in systematic ways to perfectly meet up with reality and as an AI architecture fail to scale and, therefore, need approximation. We will here discuss some of these approximations and modifications needed to reach a practical learning model. Furthermore, we will discuss some questions in philosophy of science from the perspective of this model. [Hut07,RH11] show that the classical induction problems are resolved or disappear in an ideal universal induction setting where we maintain a certain form of distribution over all computable hypotheses. This form of induction [Sol64], often called Solomonoff induction, is not only impractical but also incomputable, though still serves as an important gold standard. [Hut05] extended Solomonoff Induction to the reinforcement learning setting, which includes most (perhaps all) known sequential decision making problems.

In this work we consider a framework where we have an implicit distribution over all possible hypotheses in the form of a sampling process. This process slowly builds a class of explicit hypothesis which are excluded when they are contradicted by experience. The Bayesian learning is still present in the implicit distribution, while we will here consider a different approach for working with the explicit hypothesis. We extend our previous optimistic reinforcement learning agent presented in [SH12] to a setting with a growing hypotheses class and, furthermore, to a setting where the agent works with partial hypothesis that only make predictions for some of the features of the world. In our full setting we combine deterministic laws that predict the effect of actions on some features with correlations between features that land us in between the deterministic and stochastic settings that retains the obvious exclusion rule from the deterministic case as well as the desirable finite error bound while allowing for uncertainty in the world.

Outline. We first give an overview in Section 2. In Section 3 we introduce the algorithm for how to act given a set of theories. This is an extension of earlier work on optimistic agents for finite classes of environments to the case of partial theories and also to the case with a countably infinite number of environments. In Section 4 we define the process of generating new theories. In Section 5 we discuss the practical planning problems involved and how they relate to human cognitive biases. In Section 6 we discuss issues in philosophy of science related to our framework and we conclude in Section 7.

2 Overview of the Framework

In this section, we give an overview of the framework we suggest and its motivation before giving more technical details in the next two sections.

Restricted explicit hypothesis classes. The main restriction we impose on the agent is that it can only explicitly represent and work with a limited number of hypotheses. The hypotheses will be generated by a sampling process over time and they will be excluded when contradicted by observations. The restriction of limited hypothesis classes obviously applies to both science, where the existence of unconceived alternatives to existing hypotheses has been used to challenge the rationality of science [Sta06], and to individual human cognition where the very same issue is even more pronounced. In practical Bayesian reinforcement learning [Str00], an important technique to approximately maximize expected value is to sample a few hypothesis from the posterior and work only with these.

Having an implicit universal distribution. Here we are going to suggest a framework for learning when an unlimited amount of theories cannot be considered at a time. This model of learning can e.g. be considered for building an artificial intelligent agent or for modeling science or human cognition. It maintains an implicit generative distribution over all possible theories. In the AI setting, samples would be generated by an algorithm. Inductive Logic Programming procedures [Mug91] as well as evolutionary algorithms [Rec71,Hol75] are fields studying processes for generating hypotheses without directly aiming for sampling from a given distribution, while Markov Chain Monte Carlo (MCMC) methods , e.g. Gibbs sampling [CG92,Ber04], are directly built for the purpose of sampling from a posterior distribution. A recent article, Logical Prior Probability [Dem12], samples sentences from a first order theory for the purpose of approximating a universal distribution over sequences.

In the case of science, the scientific community and its processes and members generate new hypotheses. We model this as sampling from an implicit distribution and we stipulate that the distribution depends on both how likely a theory seems given existing evidence and how simple the theory is, possibly given the theories that have been suggested in the past. For practicality of artificial intelligent agents, and to align with reality when modeling science and cognition, we should also have a bias towards convenient theories that one can work efficiently with. As an example, note the ubiquity of tree structures in scientific theories.

Laws, concepts and correlations. The word *theory* is in cognitive science often used for a framework of concepts and relations [GT09,KTNG10]. Concepts are here understood as a feature vector that summarizes all perceptions made up until the current time. We distinguish between the correlations among the features and deterministic laws that state how some of the features change when a given action is taken. The features, the correlations and the laws make up what we here call a theory. The conceptual framework, i.e. the feature map and the correlations, allows the raw perceptions from the world to be represented by an extracted feature vector. The features could in principle be the raw data, i.e. a coding of the observed world [Hut12], but not in practise. One approach is to base the conceptual framework on a probabilistic model [KTNG10] of the world. The feature map, or rather its output, is in this paper assumed given. An example of a feature vector is the location and velocity of a number of object, e.g. planets. Furthermore, given some of the elements of the feature vector, a distribution over the rest is also assumed and referred to as correlations. The feature map and correlations must in practise be learnt, e.g. through an hierarchical Bayesian inference procedure [GT09] or as in the feature reinforcement learning framework [Hut09] by minimizing a cost function based on the ability to predict future rewards.

The laws mentioned earlier are defined on the feature vector above and are such that given the current feature vector and an action they predict some of the features at the next time step. Combined with the correlations, a probabilistic prediction about the whole next feature vector is made. A feature of particular interest is the reward received.

Reinforcement learning. A question that is often not explicitly discussed when modeling science or cognition is how decisions, e.g. choice of experiment to perform, are made. The Bayesian inference framework can be adopted to perform this job if a utility function is defined, i.e. how good is it to have a certain outcome of a chosen experiment. Given this, one chooses to act so as to maximize expected utility. This is the description of a rational agent in rational choice theory [Sav54] and in rational reinforcement learning [SH11]. The question if ideal reinforcement learning is possible and what it is, is not as clear as in the case of sequence prediction. In induction one can have most or all relevant forms of optimality, except for efficiency in computation and memory, at the same time [Hut07]. In reinforcement learning, the observation received depends on what decisions one makes, e.g. which experiments one performs. This leads to the exploration vs exploitation dilemma where one chooses between exploiting the knowledge one has or explore to gain more knowledge. However, if the goal is pure knowledge seeking, i.e. we have defined the reward to represent gain of new knowledge in a suitable sense, the dilemma disappears. In [Ors11,OL13] it is shown that a pure knowledge seeking Bayesian agent, if suitably defined, is guaranteed asymptotic optimality as well as success in identifying the correct hypothesis from all other hypotheses that are sufficiently separable from it. Asymptotic optimality cannot in general be guaranteed in reinforcement [LH11], though it is possible in a weaker "in average" sense. This failure is due to a never ending need for exploration phases. These phases do not need to be explicitly defined but one can instead rely on optimism as in [SH12] which introduced an agent that always achieves optimality for any finite class of deterministic environments and with high probability in the stochastic case. The optimistic agent introduced by [SH12] and its guarantees are in [SH13] extended to the setting of a countable class by slowly adding environments to a finite class, as well as to a setting where it has to work with partial hypothesis which can be combined and complemented with correlations to complete environments in an uncountably infinite number of ways but still satisfy error bounds depending on the number of introduced laws.

3 Optimistic Agents

We consider observations of the form of a feature vector $o = \bar{x} = (x_j)_{j=1}^{M}$ (including the reward as one coefficient) where x_j is an element of some finite alphabet. A law is a function $\tau : \mathcal{H} \times \mathcal{A} \to \tilde{\mathcal{O}}$ where $\mathcal{A}$ is the action set, $\mathcal{H}$ the set of histories and $\tilde{\mathcal{O}}$ consists of the feature vectors from $\mathcal{O}$ (the observation set) but where some elements are replaced by a special letter meaning that there is no prediction for this feature. At any time t, there is a set $\mathcal{T}_t$ of laws that have been suggested so far. We consider some of them to have been refuted since they were suggested and we have a smaller working set $\mathcal{T}_t^w \subset \mathcal{T}_t$ of unrefuted laws. The laws in the sets can be overlapping and contradictory. We combine several laws for each time step. When they cannot be combined into a complete environment, [SH13] combine laws with correlations to define complete environments.

Given a finite class of deterministic environments $\mathcal{M} = \{\nu_1, ..., \nu_m\}$, [SH12] defines an algorithm that for any unknown environment from $\mathcal{M}$ eventually achieves optimal behavior in the sense that there exists T such that maximum reward is achieved from time T onwards. The algorithm chooses an optimistic hypothesis from $\mathcal{M}$ in the sense that it picks the environment in which one can achieve the highest reward and then the policy that is optimal for this environment is followed. If this hypothesis is contradicted by the feedback from the environment, a new optimistic hypothesis is picked from the environments that are still consistent with h. This technique has the important consequence that if the hypothesis is not contradicted we are still acting optimally when optimizing for this incorrect hypothesis.

Combining laws. Each law τ predicts either none, some or all of the features (concepts) x_j at the same time point. A law depends on some of the features in the history, perhaps only on the current features if the features summarize all relevant information up to the current time. If the latter is true, it makes the task simpler because it restricts the total class of laws. Another distinction is whether all features are such that they can be predicted (near) deterministically from the previous features or not. Given a set of laws where in every situation and for every feature there is at least one law that makes a prediction of this feature in the given situation, then we can directly define a set of environments by combining such laws. If this is not the case, we will need to use the relations between the features. [SH13] extends the optimistic agent from [SH12] to this case by only considering contradiction for the deterministically predicted features. Whenever an environment is contradicted, some law must be contradicted. Therefore, the number of introduced laws up to a certain time upper bounds how many contradictions we could have had. After the true environment has been introduced this bounds the number of errors committed by the agent. Since the true environment is introduced after a finite amount of time, this reasoning, but more formally defined, leads to the error bounds in [SH13].

Example 1 (Going to get coffee). If the agent from a state where it is in its office (a boolean feature) goes to the coffee machine in the kitchen (a sequence of actions) it ends up at the coffee machine (a sequence of laws). Standing at the coffee machine correlates to some probabilistic degree with seeing that there is too many people being at the coffee machine. Furthermore, if it is after five on a Friday and then the agent will end up in a state where it is at the coffee machine and it is after five on a Friday. Conditioning on these two features, the probability for the coffee machine being crowded and no reward of coffee, is lower. These deterministic laws have to be learnt and an agent might start with many more alternatives for what happens when going into the kitchen, e.g. ending up on the street. Exploration rules out wrong alternatives.

Joint learning of concepts and laws. In reality, this article's assumption that we have already finished learning the conceptual framework (features and correlations) and that it needs no further updating, is not realistic. Also, since there might be some arbitrariness in representation of the distribution over raw data that the conceptual framework is based on, we want a representation with features that admit powerful deterministic laws. This leads to an interplay between learning laws and concepts. A particularly interesting situation is when a suggested law is contradicted. Then, this might either lead to excluding the law or modifying the involved concepts.

4 Suggesting Laws

We take the position that a law's probability of being suggested is a combination of the probability that it will be considered based on simplicity (possibly given previous suggestions) and the probability that it is true given what we know. In other words, the law's appearance or not depends probabilistically on both the simplicity of the theory and the alignment with observations made. We will assume that these two are independent, i.e.

$$Pr(\tau \textit{ suggested } |h_t) = Pr(\tau \textit{ considered})Pr(\tau \text{ true}|h_t)$$

or if we take previously suggested theories into account

$$Pr(\tau \text{ suggested}|h_t, \mathcal{T}_t) = Pr(\tau \text{ considered}|\mathcal{T}_t)Pr(\tau \text{ true}|h_t).$$

In [GS82,HLNU13] it was shown that one can define distributions over rich languages like first order logic or even higher order logic and that the distributions satisfy properties that makes them suitable for learning, i.e. for defining $Pr(\tau \text{ true}|h_t)$. For example, there are distributions that satisfy Cornout's condition that if something is (logically) possible it must have strictly positive probability and the Gaifman condition that enables confirmation of universal laws in the limit.

To define a sensible distribution to sample laws from, we rely on Algorithmic Information Theory and the notion of Kolmogorov complexity [LV93]. The resulting distributions have the property that laws that are simpler, in the sense of being compressible into a short code, are more likely. Sampling from such a distribution is what we consider to be ideal, though we recognize that any practical generative process is likely to have a sequential dependence where new laws that are similar to already suggested laws are more likely to be considered though not necessarily suggested. Therefore, after some laws (sentences) have been suggested, we need to consider laws that are close to those earlier ones, to be simpler than they were because we can now understand them from the perspective of a law we already know. This is naturally formalized using conditional Kolmogorov complexity $K(s|s')$ which is the length of the shortest program that produces output (the encoding of) s given side information s'. If sentences $s_1, ..., s_n$ have been suggested, we let

$$Pr(\tau \textit{ considered}|s_1, ..., s_n) = 2^{-K(\tau|s_1,...,s_n)}.$$

Example 2. If τ has been suggested and τ says that if $x_j = a$ at time t then $x_k = b$ at time $t+1$, then the law that instead concludes that $x_k = c \neq b$ is more likely to be considered but it might still not be likely to be suggested if the data is aligned with the predictions of τ.

5 Human Cognitive Biases and Practical Planning

We have so far ignored the issue of how to calculate the optimal value that can be achieved in some environment. The task of finding the optimal policy, i.e the policy that achieves the highest expected value, in some given environment is the planning problem [RN10]. This is a very demanding problem and much work in AI on planning has been conducted in a deterministic setting which greatly reduces the complexity of the task. The optimistic choice can actually be represented as just one planning problem but with the action space enlarged to include the choice of laws to use for predicting the next step. This trick has been used in the Markov Decision Process case [ALL$^+$09].

Science is always aiming at finding concepts which allow for deterministic prediction and the hind-sight bias [NBvC08] in human cognition, referring to the human habit of finding a way of seeing past events as inevitable, can be viewed as searching for deterministic explanations.

Optimism treats the uncertainty regarding which laws are true in a positive way where we assume that we are in the best of all possible worlds. This is a strong version of the optimism bias in human cognition [Wei80]. As we have discussed earlier, optimism leads to explorative behavior though obviously this best-of-all-worlds assumption is also associated with big risks. Humans incorporate a strong aversion to big losses in their decision making [KT79].

6 Philosophical Issues

In this section we discuss how one can understand a number of issues studied in philosophy of science assuming our model of learning.

Paradigms, unconceived alternatives and rationality of science.
[Sta06] discussed the problem of believing in a hypothesis because one cannot conceive of alternatives and asks if this makes science irrational. From the strict point of view that the rational choice axioms [Sav54] lead to Bayesian inference where we have a prior, which should be strictly positive for all alternatives, to be admissible, it is correct to conclude that science is not rational since it is only working with a subset. However, we argue that it is enough that the processes of science will *eventually* produce those unconceived alternatives and that the likelihood of them appearing will increase when evidence makes them more likely to be true. This approach saves a claim to a weaker form of rationality where we demand; a) that a hypothesis' likelihood of appearing depends on its likelihood to be true in the Bayesian sense of conditional probability given a prior distribution over all possible hypothesis; b) When it has appeared it is evaluated based on if it aligns with evidence or not; c) Experiments are chosen such that eventually, if not immediately, we will falsify wrong hypothesis.

A consequence of the discussed model is that wrong hypotheses might be retained for a long time if they are useful and not contradicted but eventually better alternatives are conceived of and the old hypothesis is overturned. In the meantime we are acting in a manner that is confirming the existing fruitful hypotheses. This process naturally leads to paradigms as discussed by [Kuh70].

Old evidence and new theories. The "old evidence and new theories problem" asks how evidence gathered before a theory was suggested can be used, as it has often been, to validate the new theory. In the universal Bayesian approach all possible theories (e.g. all computable hypotheses) are included from the start and, therefore, there are no new theories and no problem [Hut07,RH11]. However, this is not the situation of actual science where the conundrum came from. The answer we present here is in a sense close to the answers from [Hut07,RH11] though we have a limited explicit set of theories generated through a process that (approximately) samples from an implicit distribution over all possible theories with a bias towards simple theories. It is this full distribution that is being updated through the processes of the scientific community and this resolves the problem without removing the question. A combination of a theory's simplicity and its alignment with past observations makes the theory plausible and warrants inclusion among the small number of theories whose predictions are being further tested.

Theory-laden observations and reintroduced hypotheses. In the history of science, excluded theories have sometimes been reintroduced. This can happen if the scientific community due to new evidence that strongly supports a previously rejected hypothesis, suspects that the previous observations were not valid or misinterpreted. The latter can relate to a modification of the conceptual framework that underpins the interpretation of raw data. The observations depend on the conceptual framework, which is often expressed as saying that they are theory-laden [Han58]. This is what we discuss in Section 3 when addressing joint learning of concepts and laws.

7 Conclusions

Bayesian inference is often viewed as ideal learning in many different fields including AI, cognitive science and philosophy of science. However, many questions in those fields do not make sense in the full Bayesian framework and when creating artificial intelligent agents, approximations to scale to complex environments are required. In all of these areas it is clear that one cannot explicitly maintain beliefs over truly rich sets of hypotheses. We introduced a framework for learning based on implicit beliefs over all possible hypotheses and limited sets of explicit theories. Furthermore, we showed how to define a reasonable instance of the framework and we discussed implications for some questions in philosophy of science.

Acknowledgement. This work was supported by ARC grant DP120100950.

References

[ALL+09] John Asmuth, Lihong Li, Michael L. Littman, Ali Nouri, and David Wingate. A bayesian sampling approach to exploration in reinforcement learning. In *Uncertainty in Artificial Intelligence (UAI)*, pages 19–26, 2009.

[And91] John R. Anderson. Is human cognition adaptive? *Behavioral & Brain Sciences*, 14(3):471–517, 1991.

[Ber04] Bernd A. Berg. *Markov Chain Monte Carlo Simulations And Their Statistical Analysis: With Web-based Fortran Code.* World Scientific Publishing Company, 2004.

[BT92] George E. P. Box and George C. Tiao. *Bayesian Inference in Statistical Analysis (Wiley Classics Library).* Wiley-Interscience, 1992.

[CG92] George Casella and Edward I. George. Explaining the gibbs sampler. *The American Statistician*, 46(3):167–174, 1992.

[Dem12] Abram Demski. Logical prior probability. In Joscha Bach, Ben Goertzel, and Matthew Iklé, editors, *AGI*, volume 7716 of *Lecture Notes in Computer Science*, pages 50–59. Springer, 2012.

[Edw68] W. Edwards. Conservatism in human information processing. *Formal representation of human judgment*, pages 17–52, 1968.

[GKT08] Thomas L. Griffiths, Charles Kemp, and Joshua B. Tenenbaum. *Bayesian Models of Cognition.* Cambridge University Press, 2008.

[Goo84] I. J. Good. A Bayesian approach in the philosophy of inference. *British Journal for the Philosophy of Science*, pages 161–166, 1984.

[GS82] Haim Gaifman and Marc Snir. Probabilities over rich languages, testing and randomness. *J. Symb. Log.*, 47(3):495–548, 1982.

[GT09] Thomas L. Griffiths and Joshua B. Tenenbaum. Theory-based causal induction. *Psychological Review*, 116(4):661–716, 2009.

[Han58] Norwood Russell Hanson. *Patterns of Discovery.* Cambridge [Eng.]University Press, 1958.

[HLNU13] M. Hutter, J.W. Lloyd, K.S. Ng, and W.T.B. Uther. Probabilities on sentences in an expressive logic. *Journal of Applied Probability*, 2013.

[Hol75] John H. Holland. *Adaptation in Natural and Artificial Systems.* University of Michigan Press, 1975.

[HU05] Colin Howson and Peter Urbach. *Scientific Reasoning: The Bayesian Approach.* Open Court, third edition edition, 2005.

[Hut05] M. Hutter. *Universal Articial Intelligence: Sequential Decisions based on Algorithmic Probability.* Springer, Berlin, 2005.

[Hut07] Marcus Hutter. On universal prediction and bayesian confirmation. *Theoretical Computer Science*, 384:33–48, 2007.

[Hut09] M. Hutter. Feature reinforcement learning: Part I. Unstructured MDPs. *Journal of General Artificial Intelligence*, 2009.

[Hut12] Marcus Hutter. The subjective computable universe. In *A Computable Universe: Understanding and Exploring Nature as Computation*, pages 399–416. World Scientific, 2012.

[Jay03] E. T. Jaynes. *Probability Theory: The Logic of Science.* Cambridge University Press, 2003.

[KT79] D. Kahneman and A. Tversky. Prospect theory: An analysis of decision under risk. *Econometrica*, 47:263–291, 1979.

[KTNG10] C. Kemp, J. B. Tenenbaum, S. Niyogi, and T. L. Griffiths. A probabilistic model of theory formation. *Cognition*, 114(2), 2010.

[Kuh70] Thomas S. Kuhn. *The structure of scientific revolutions.* University of Chicago Press, 1970.

[LH11] T. Lattimore and M. Hutter. Asymptotically optimal agents. In *Proc. of Algorithmic Learning Theory (ALT'2011)*, volume 6925 of *Lecture Notes in Computer Science*, pages 368–382. Springer, 2011.

[LV93] M. Li and P. Vitany. *An Introduction to Kolmogov Complexity and Its Applications.* Springer, 1993.

[Mug91] Stephen Muggleton. Inductive logic programming. *New Generation Computing*, 8(4):295–318, 1991.

[NBvC08] S. Nestler, H. Blank, and G. von Collani. Hindsight bias doesn't always come easy: Causal models, cognitive effort, and creeping determinism. *Journal of Experimental Psychology: Learning*, 34(5), 2008.

[Nya93] Y. Nyarko. The savage-bayesian foundations of economic dynamics. Working Papers 93-35, C.V. Starr Center for Applied Economics, New York University, 1993.

[OL13] L. Orseau and T. Lattimore. Univeral knowledge-seeking agents for stochastic environments. submitted 2013.

[OR94] Martin J. Osborne and Ariel Rubinstein. *A Course in Game Theory.* MIT Press Books. The MIT Press, 1994.

[Ors11] Laurent Orseau. Universal knowledge-seeking agents. In *Algorithmic Learning Theory - 22nd International Conference, (ALT 2011)*, volume 6925 of *Lecture Notes in Computer Science*, pages 353–367. Springer, 2011.

[Rec71] I. Rechenberg. *Evolutionsstrategie.* Frommann-Holzboog-Verlag, 1971.

[RH11] Samuel Rathmanner and Marcus Hutter. A philosophical treatise of universal induction. *Entropy*, 13(6):1076–1136, 2011.

[RN10] S. J. Russell and P. Norvig. *Artificial Intelligence: A Modern Approach.* Prentice Hall, Englewood Cliffs, NJ, 3^{nd} edition, 2010.

[Rob94] Christian P. Robert. *The Bayesian choice : a decision-theoretic motivation / Christian P. Robert.* Springer-Verlag, New York :, 1994.

[Sav54] L. Savage. *The Foundations of Statistics.* Wiley, New York, 1954.

[SH11] P. Sunehag and M. Hutter. Axioms for rational reinforcement learning. In *Algorithmic Learning Theory (ALT'2011)*, volume 6925 of *Lecture Notes in Computer Science*, pages 338–352. Springer, 2011.

[SH12] P. Sunehag and M. Hutter. Optimistic agents are asymptotically optimal. In *Proceedings of the 25:th Australasian AI conference*, pages 15–26, 2012.

[SH13] P. Sunehag and M. Hutter. Rational general reinforcement learning. submitted 2013.

[Sol64] Ray J. Solomonoff. A formal theory of inductive inference. part i and ii. *Information and Control*, 7(1,2):1–22,224–254, 1964.

[Sta06] P. Kyle Stanford. *Exceeding Our Grasp: Science, History, and the Problem of Unconceived Alternatives.* Oxford University Press, 2006.

[Str00] Malcolm J. A. Strens. A bayesian framework for reinforcement learning. In *Proceedings of the Seventeenth International Conference on Machine Learning (ICML'2000)*, pages 943–950, 2000.

[Str06] Michael Strevens. The bayesian approach to the philosophy of science. In *Encyclopedia of Philosophy, second edition.* 2006.

[TKGG11] Joshua B. Tenenbaum, Charles Kemp, Thomas L. Griffiths, and Noah D. Goodman. How to Grow a Mind: Statistics, Structure, and Abstraction. *Science*, 331(6022):1279–1285, 2011.

[Wei80] Neil D. Weinstein. Unrealistic optimism about future life events. *Journal of Personality and Social Psychology*, 39(5):806–820, 1980.