

Experience-Oriented Artificial Intelligence

Abstract

AI is at an impasse. It is stuck, or downsizing. Unable to build large, ambitious systems because no means to manage complexity. Now people manage complexity, but a large AI must do it itself. An AI must be able to tell for itself when it is right and when it is wrong. Experience is the route to this...

Experience should be at the center of AI. It is what AI is about. It is the data of AI, yet it has been sidelined. An AI must be able to tell for itself when it is right and when it is wrong.

Experience plays a central role in the problem of artificial intelligence. If intelligence is a computation, then the temporal stream of sensations is its input, and the temporal stream of actions is its output. These two intermingled time series are both the basis for all intelligent decision making and the basis for assessing it. Experience waits for neither man nor machine. Its events occur in an unalterable order and pace. Sensory signals may require quick action, or a more deliberate response. An action taken cannot be retracted. the temporal structure of experience may be the single most important computational feature of the problem of artificial intelligence.

Nevertheless, experience has played a less than salient role in the *field* of artificial intelligence. Artificial intelligence has often dealt with subjects such as inference, diagnosis and problem-solving in such a way as to minimize the impact of real-time sensation and action. It is hard to discern any meaningful role for experience in classical question-answering AI systems. These systems may help people predict and control *their* experience, but the systems themselves have none.

Robotics has always been an important exception, but even there experience and time play less of a role than might have been anticipated. Motor control is dominated by planning methods that emphasize trajectories and kinematics over dynamics. Computer vision research is concerned mostly with static images, or with open-loop streams of images with little role for action. Machine learning is dominated by methods which assume independent, identically distributed data—data in which order is irrelevant and there is no action.

Recent trends in artificial intelligence can be seen as in part a shift in orientation towards experience. The “agent oriented” view of AI can be viewed in this light. Probabilistic models such as Markov decision processes, dynamic Bayes networks, and reinforcement learning are also part of the modern trend towards recognizing a primary role for temporal data and action.

A natural place to begin exploring the role of experience in artificial intelligence is in knowledge representation. Knowledge is critical to the performance of successful AI systems, from the knowledgebase of a diagnosis system to the evaluation function of a chess-playing program to the map and sensor model of a navigating robot. Intelligence itself can be defined as the ability to maintain a very large body of knowledge and apply it effectively and flexibly to new problems.

While large amounts of knowledge is a great strength of AI systems, it is also a great weakness. The problem is that as knowledge bases grow they become more brittle and difficult to maintain. There arise inconsistencies in the terminology used by different people or at different times. The more diverse the knowledge the greater are the opportunities for confusions. Errors are inevitably present, if only because of typos in data entry. When an error becomes apparent, the problem can only be fixed by a human who is expert in the structure and terminology of the knowledge base. This is the root difficulty: the accuracy of the knowledge can ultimately only be verified and safely by a person intimately familiar with most of the knowledge and its representation. This puts an upper bound on the size of the knowledge base. As long as people are the ultimate guarantors—nay, definers—of truth, then the machine cannot become much smarter than its human handlers. Verifying knowledge by consistency with human knowledge is ultimately, inevitably, a dead end.

How can we move beyond human verification? There may be several paths towards giving the machine more responsibility and ability for verifying its knowledge. One is to focus on the consistency of the knowledge. It may be possible to rule out some beliefs as being logically or mathematically inconsistent. For the vast majority of everyday world knowledge, however, it seems unlikely that logic alone can establish truth values.

Another route to verification, the one explored in this paper, is consistency with experience. If knowledge is expressed as a statement about experience, then in many cases it can be verified by comparison with experiential data. This approach has the potential to substantially resolve the problem of autonomous knowledge verification. [some examples: battery charger, chair, john is in the coffee room] The greatest challenge to this approach, at least as compared with human verification, is that sensations and actions are typically low-level representations, whereas the knowledge that people most easily relate to is at a much higher level. This mismatch makes it difficult for people to transfer their knowledge in an experiential form, to understand the AI's decision process, and to trust its choices. But an even greater challenge is to our imaginations. How is it possible for even slightly abstract concepts, such as that of a book or a chair, to be represented in experiential terms? How can they be represented so fully that everything about that concept has been captured and can be autonomously verified? This paper is about trying to answer this question.

First I establish the problem of experiential representation of abstract concepts more formally and fully. That done, an argument is made that all world knowledge is well understood as predictions of future experience. Although the gap from low-level experience to abstract concepts may seem immense, in theory it must be bridgeable. The bulk of this paper is an argument that this which in theory must be, is also plausible. Recent methods for state and action representation, together with function approximation, can enable us to take significant steps toward abstract concepts that are fully grounded in experience.

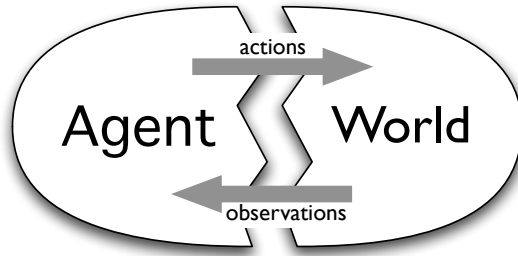


Figure 1: Experience is the signals crossing the line separating agent from world.

1. Experience

To distinguish an agent from its world is to draw a line. On one side is the agent, receiving sensory signals and generating actions. On the other side, the world receives the actions and generates the sensory signals. Let us denote the action taken at time t as $a_t \in \mathcal{A}$, and the sensation, or *observation*, generated at time t as $o_t \in \mathcal{O}$. Time is taken to be discrete, $t = 1, 2, 3, \dots$. The time step could be arbitrary in duration, but we think of it as some fast time scale, perhaps one hundredth or one thousandth of a second. Experience is the intermingled sequence of actions and observations

$$o_1, a_1, o_2, a_2, o_3, a_3, \dots$$

each element of which depends only on those preceding it. See Figure 1. Define $\mathcal{E} = \{\mathcal{O} \times \mathcal{A}\}^*$ as the set of all possible experiences.

Let us call the experience sequence up through some action a *history*. Formally, any world can be completely specified by a probability distribution over next observations conditional on history, that is, by the probability $P(o|h)$ that the next observation is o given history h , for all $o \in \mathcal{O}$ and $h \in \mathcal{E}$. To know P exactly and completely is to know everything there is to know about the agent's world.¹ Short of that, we may have an approximate model of the world.

2. Predictive knowledge

The perspective being developed here is that the world is a formal, mathematical object, a function mapping histories to probability distributions over observations. In this sense it is pointless to talk about what is “really” going on in the world. The only thing to say about the world is to predict probability distributions over observations. This is meant to be an absolute statement. Given an input-output definition of the world, there can be no knowledge of it that is not experiential:

1. Suppose we have a *model* of the world, an approximation \hat{P} to P . How can we define the quality of the model? First, we need only look at the future; we can take the history so far as given and just consider further histories after that. Thus, \hat{P} and P can be taken to give distributions for future histories. I offer a policy-dependent measure of the *loss* of a model, that is, of how much it does not predict the data:

$$L_\pi(P||\hat{P}) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{l=0}^n \frac{1}{|H_l|} \sum_{h \in H_l} \sum_o P(o|h) \log \frac{1}{\hat{P}(o|h)}.$$

Everything we know that is specific to this world (as opposed to universally true in any world) is a prediction of experience. All world knowledge must be translatable into statements about future experience.

Our focus is appropriately on the predictive aspect. Memories can be simple recordings of the full experience stream to date. Summaries and abstract representations of the history are significant only in so far as they affect predictions of future experience. Without loss of generality we can consider all world knowledge to be predictive.

One possible objection could be that logical and mathematical knowledge is not predictive. We know that $1+1 = 2$, or that the area of a circle is πr^2 , or that $\neg(p \vee q) \Leftrightarrow \neg p \wedge \neg q$, and we know these absolutely. Comparing them to experience cannot prove them wrong, only that they do not apply in this situation. Mathematical truths are true for any world. However, for this very reason they cannot be considered knowledge of any particular world. Knowing them may be helpful to us as part of making predictions, but only the predictions themselves can be considered world knowledge.

These distinctions are well known in philosophy, particularly the philosophy of science. Knowledge is conventionally divided into the analytic (mathematical) and the synthetic (empirical). The logical positivists were among the earliest and clearest exponents of this point of view and, though it remains unsettled in philosophy, it is unchallenged in science and mathematics. In retrospect, mathematical and empirical truth—logical implication and accurate prediction—are very different things. It is unfortunate that the same term, “truth,” has been used for both.

Let us consider some examples. Clearly, much everyday knowledge is predictive. To know that Joe is in the coffee room is to predict that you will see him if you went there, or that you will hear him if you telephone there. To know what’s in a box is to predict what you will see if you open it, or hear if you shake it, feel if you lift it, and so on. To know about gravity is to make predictions about how objects behave when dropped. To know the three-dimensional shape of an object in your hand, say a teacup, is to predict how its silhouette would change if you were to rotate it along various axes. A teacup is not a single prediction but a pattern of interaction, a coherent set of relationships between action and observation.

Other examples: Dallas Cowboys move to Miami. My name is Richard. Very cold on pluto. Brutus killed Caesar. Dinosaurs once ruled the earth. Canberra is the capital of Australia. Santa Claus wears a red coat. A unicorn has one horn. John loves Mary.

Although the semantics of “Joe is in the coffee room” may be predictive in an informal sense, it stills seems far removed from an explicit statement about experience, about the hundred-times-a-second stream of inter-mingled observations and actions. What does it mean to “go to the coffee room” and “see him there”. The gap between everyday concepts and low-level experience is immense. And yet there must be a way to bridge it. The only thing to say about the world is to make predictions about its behavior. In a formal sense, anything we know or could know about the world must be translatable into statements about low-level future experience. Bridging the gap is a tremendous challenge, and in this paper I attempt to take the first few steps toward it. This is what I call the grand challenge of grounding knowledge in experience:

To represent human-level world knowledge solely in terms of experience, that is, in terms of observations, actions, and time steps, without reference to any other concepts or entities unless they are themselves represented in terms of experience.

The grand challenge is to represent all world knowledge with an extreme, minimalist ontology of only three elements. You are not allowed to presume the existence of self, of objects, of space, of situations, even of “things”.

Grounding knowledge in experience is extremely challenging, but brings an equally extreme benefit. Representing knowledge in terms of experience enables it to be *compared with* experience. Received knowledge can be verified or disproved by this comparison. Existing knowledge can be tuned and new knowledge can be created (learned). The overall effect is that the AI agent may be able to take much more responsibility for maintaining and organizing its knowledge. This is a substantial benefit; the lack of such an ability is obstructing much AI research, as discussed earlier.

A related advantage is that grounded knowledge may be more useful. The primary use for knowledge is to aid planning or reasoning processes. Predictive knowledge is suited to planning processes based on repeated projection, such as state-space search, dynamic programming (), and model-based reinforcement learning (Dyna, pri-sweep, LSTD). If A predicts B, and B predicts C, then it follows that A predicts C. If the final goal is to obtain some identified observation or observations, such as rewards, then predictive reasoning processes are generally suitable.

3. Questions and Answers

Modern philosophy of science tells us that any scientific theory must be empirically verifiable. It must make predictions about experiments that can be compared to measurable outcomes. We have been developing a similar view of knowledge—that the content of knowledge is a prediction about the measurable outcome of a way of behaving. The prediction can be divided into two parts, one specifying the question being asked and the other the answer offered by the prediction. The question is “What will be the measured value if I behaved this way and measured that?” An answer is a particular predicted value for the measurement which will be compared to what actually happens to assess the prediction’s accuracy. For example, a question roughly corresponding to “How much will it rain tomorrow” would be a procedure for waiting, identifying when tomorrow has begun, measuring the cumulative precipitation in centimeters, and ending when the end-of-day has been identified. The result based on this actual future will be a number such as 1.2 which can be compared to the answer offered by the prediction, say 1.1.

In this example, the future produces a result, the number 1.2, whose structure is similar to that of the answer, and one may be tempted to refer to the result as the “correct answer.” In general, however, there will be no identifiable correct answer that can be identified as arising from the question applied to the future. The idea of a correct answer is also misleading because it suggests an answer coming from the future, whereas we will consider answers always to be generated by histories. There may be one or more senses of best answers that could be generated, but always from a history, not a future.

Figure 3 shows how information flows between experience and the question and answer making up a prediction made at a particular time. Based on the history, the answer is formed and passed

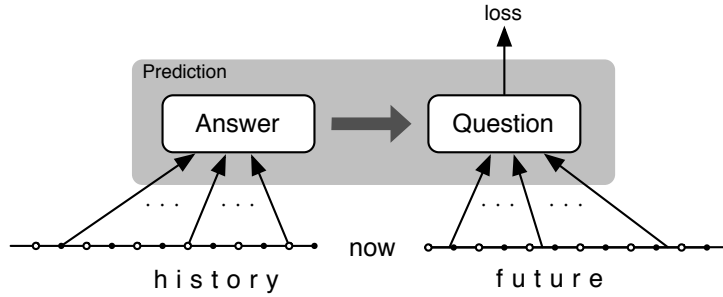


Figure 2: Information flow relationships between questions and answers, histories and futures.

on to the question, which compares it with the future. Eventually, a measure of mismatch between answer and future is computed, called the loss. This process is repeated at each moment in time and for each prediction made at that time.

Note that the question in this example is substantially more complex and substantial than its answer; this is typically the case. Note also that the question alone is not world knowledge. It does not say anything about the future unless matched with an answer.

For knowledge to be clear, the experiment and the measurement corresponding to the question must be specified unambiguously and in detail. We state this viewpoint as the *explicit prediction manifesto*:

Every prediction is a question and an answer.

Both the question and the answer must be explicit in the sense of being accessible to the AI agent, i.e., of being machine readable, interpretable, and usable.

The explicit prediction manifesto is a way of expressing the grand challenge of empirical knowledge representation in terms of questions and answers. If knowledge is in predictive form, then the predictions must be explicit in terms of observations and actions in order to meet the challenge.

It is useful to be more formal at this point. In general, a question is a loss function on futures with respect to a particular way of behaving. The way of behaving is formalized as a policy, a (possibly deterministic) mapping from $\mathcal{E} \times \mathcal{O}$ to probabilities of taking each action in \mathcal{A} . The policy and the world together determine a future or probability distribution over futures. For a given space of possible answers \mathcal{Y} , a question’s loss function is a map $q : \mathcal{E} \times \mathcal{Y} \mapsto \mathbb{R}^+$ from futures and answers to a non-negative number, the *loss*. A good answer is one with a small loss or small expected loss.

For example, in the example given above for “How much will it rain tomorrow”, the answer space is the non-negative real numbers, $\mathcal{Y} = \mathbb{R}^+$. Given a history $h \in \mathcal{E}$, an answer $y(h)$ might be produced by a learned answer function $y : \mathcal{E} \mapsto \mathcal{Y}$. Given a future $f \in \mathcal{E}$, the loss function would examine it in detail to determine the time steps at which tomorrow is said to begin and end. Suppose the precipitation on each time step “in centimeters” is one component of the observation on that step. This component is summed between the start and end times to produce a correct

answer $z(f) \in \mathcal{E}$. Finally, $y(h)$ and $z(f)$ are compared to obtain, for example, a squared loss $q(f, y(h)) = (z(f) - y(h))^2$.

The interpretation in terms of “centimeters” in this example is purely for our benefit; the meaning of the answer is with respect to the measurement made by the question, irrespective of whatever interpretation we might place on it. Our approach is unusual in this respect. Usually in statistics and machine learning the focus is on *calibrated* measurements that accurately mirror some quantity that is meaningful to people. Here we focus on the meaning of the answer that has been explicitly and operationally defined by the question’s loss function. By accepting the mechanical interpretation as primary we become able to verify and maintain the accuracy of answers autonomously without human intervention.

A related way in which our approach is distinctive is that we will typically consider many questions and a great variety of questions. For example, to express the shape of an object alone requires many questions corresponding to all the ways the object can be turned and manipulated. In statistics and machine learning, on the other hand, it is common to consider only a single question. There may be a training set of inputs and outputs with no temporal structure, in which case the single question “what is the output for this input?” is so obvious that it needs little attention. Or there may be a time sequence but only a single question, such as “what will the next observation be?”

In these cases, in which there is only one questions, it is common to use the word “prediction” to refer just to answers. In machine learning, for example, the output of the learner—the answer—is often referred to as a prediction. It is important to realize that that sense of prediction—without the question—is much smaller than that which we are using here. Omitting the question is omitting much; the question part of a prediction is usually much larger and more complex than the answer part. For example, consider the question, “If I flipped this coin, with what probability would it come up heads?” The answer is simple; it’s a number, say 0.5, and it is straightforward to represent it in a machine. But how is the machine to represent the concepts of *flipping*, *coin*, and *heads*? Each of these are high-level abstractions corresponding to complex patterns of behavior and experience. Flipping is a complex, closed-loop motor procedure for balancing the coin on a finger, striking it with your thumb, then catching, turning, and slapping it onto the back of your hand. The meaning of “heads” is also a part of the question and is also complex. Heads is not an observation—a coin showing heads can look very different at different angles, distances, lightings and positions. We will treat this issue of abstraction later in the paper, but for now note that it must all be handled within the question, not the answer. Questions are complex, subtle things. They are the most important part of a prediction and selecting which ones to answer is one of the most important skills for an intelligent agent.

All that having been said, it is also important to note that predictive questions can also be simple. Perhaps the simplest question is “what will the next observation be,” (say with a cross-entropy loss measure). Or one might ask whether the third observation from now will be within some subset. If the observations are boolean we might ask whether the logical AND of the next two will be true. If they are numeric we might ask whether the square root of the sum of the next seven will be greater than 10, or whether the sum up to the next negative observation is greater than 100. Or one can ask simple questions about action dependencies. For example, we might ask what the next observation will be given that we take a particular next action, or a particular sequence of actions. In classical

predictive state representations, the questions considered, called *tests*, ask for the probability that the next few observations will take particular values if the next few actions were to have particular values. Many of these questions (but not the last one) are meant as policy dependent. For example, if a question asks which of two observations will occur first, say death and taxes, then the answer may well depend on the policy for taking subsequent actions. These simple questions have in common that we can all see that they are well defined in terms of our minimal ontology—observations, actions, and time steps. We can also see how their complexity can be increased incrementally. The grand challenge asks how far this can be taken. Can a comparable clarity of grounding be attained for much larger, more abstract, and more complex concepts?

4. Abstract Concepts and Causal Variables

Questions and answers provide a formal language for addressing the grand challenge of grounding knowledge in experience, but do not in themselves directly address the greatest component challenge, that of abstracting from the particularities of low-level experience to human-level knowledge. Let us examine in detail a few steps from low-level experience to more abstract concepts. The first step might be to group together all situations that share the same observation. The term “situation” here must be further broken down because it is not one of our primitive concepts (observations, actions, or time steps). It must be reduced to these in order to be well-defined. What is meant by “situations” here is essentially *time steps*, as in all the time steps that share the same observation. With this definition, the concept of all such time steps is clear and explicit.

A further step toward abstraction is to define subsets of observations and group together all time steps with observations within the same subset. This is natural when observations have multiple components and the subsets are those observations with the same value for one of the components. Proceeding along the same lines, we can discuss situations with the same action, with the same action-observation combination, with the same recent history of observations and actions, or that fall within any subset of these. All of these might be called *history-based* concepts. The general case is to consider arbitrary sets of histories, $C \subset \{\mathcal{O} \times \mathcal{A}\}^*$. We define *abstract* history-based concepts to be sets such that $|C| = \infty$.

It is useful to generalize the idea of history-based concepts to that of *causal variables*—time sequences whose values depend only on preceding events. (A history-based concept corresponds to a binary causal variable.) Formally, the values of causal variable $v_t = v(h_t)$ are given by a (possibly stochastic) function $v : \mathcal{E} \mapsto \mathcal{Y}$. As with concepts, we consider a causal variable to be abstract if and only if its value corresponds to an infinite set of possible histories. Formally, we define a causal variable to be *abstract* if and only if the preimage of every subset of \mathcal{Y} is infinite ($\forall C \subseteq \mathcal{Y}, |\{e : v(e) \in C\}| = \infty$). One example of a causal variable is the time sequence of answers given by the answer function of a prediction. In this sense, answers are causal variables.

Abstract causal variables seem adequate and satisfactory to capture much of what we mean by abstractions. They capture the idea of representing situations in a variety of ways exposing potentially relevant similarities between time steps. They formally characterize the space of all abstract concepts. But it is not enough to just have abstractions; they must be good abstractions. The key remaining challenge is to identify or find abstract causal variables that are likely to be useful.

In this paper we pursue the hypothesis that non-redundant answers to predictive questions are likely to be useful abstractions. This hypothesis was first stated and tested by Rafols, Ring, Sutton, and Tanner (2005) in the context of predictive state representations. They stated it this way:

“The *predictive representations hypothesis* holds that particularly good generalization will result from representing the state of the world in terms of predictions about possible future experience.”

This hypothesis is plausible if we take the ultimate goal to be to represent knowledge predictively. The hypothesis is not circular because there are multiple questions. The hypothesis is that the answer to one question might be a particularly good abstraction for answering a second question. An abstraction’s utility for one set of questions can perhaps act as a form of cross validation for its likely utility for other questions. If a representation would have generalized well in one context, then perhaps it will in another.

The hypothesis that answers to predictive questions are likely to make good abstractions begs the question of where the predictive questions come from. Fortunately, guidance as to likely pertinent questions is available from several directions. First, predictions are generally with respect to some causal variable of interest. Interesting causal variables include:

1. Signals of intrinsic interest such as rewards, loud sounds, bright lights; signals that have been explicitly designated by evolution or designer as salient and likely to be important to the agent
2. Signals that have been found to be associated with, or predictive of, signals already identified as being of interest (e.g., those of intrinsic salience mentioned in #1)
3. Signals that can be predicted, that repay attempts to predict them with some increase in predictive accuracy, as opposed to say, random signals
4. Signals that enable the control of other signals, particularly those identified as being of interest according to #1–#3

There is a fifth property² making of a causal variable interesting as a target for prediction that is more subtle and directly relevant to the development in this paper: the causal variable may itself be an answer to a predictive question. In other words, “what will be the value of this abstraction (causal variable) in the future (given some way of behaving). Questions about abstractions known to be useful would be considered particularly appealing.

The proposal, overall, is that useful abstractions for answering predictive questions can be found as answers to other predictive questions about useful abstractions. This is not circular reasoning, but rather an important form of compositionality: the ability to build new abstractions out of existing ones. It is a key property necessary for powerful representations of world knowledge.

If questions are to be about (future values of) abstractions, then what should those questions be? Recall that questions are conditional on a way of behaving – an experiment or policy. But which

2. naah, should have done this first, then the list. Actually, probably all of it should probably be reworked to be clearer and simpler. But let us go on anyway...

experiment? Guidance comes from how the predictions will be used, which will generally be as part of a planning (optimal decision making) process. Accordingly, we are particularly interested in questions about causal variables conditional on a way of behaving that optimizes the causal variables. The terminations of experiments can be selected in the same way. To explore these issues more freely we need to look at the planning process.

5. Planning

Here we discuss planning, planning with options, and the overall idea of options, including option models as predictions. This will require generalizations of options and predictions. Options will be generalized to use recognizers and option models will be generalized to be about arbitrary causal variables rather than states. The notion of predictions will have to be extended to include bundles of predictions that share a recognizer and perhaps a termination criterion but differ in the causal variables predicted.

The planning method could be "reinforcement planning," meaning planning by generating hypothetical experience and then learning from it.

6. Learning algorithm

Here we present a new TD-network or UPK algorithm. Call it UPK 2 for now.

action-filter c observation-filter beta loss?

7. An illustration

Put all the ideas together. Rat in a maze. this is an example where the abstractions convert from the low, low level of time observations and actions to high-level ideas of corridors and choice points. We do not attempt to learn this, but only to build them in with some automatic tuning. The point is that the abstractions will really be there, not just glossed over with verbal descriptions. The relationship between the abstractions and the low-level stuff will be explicit and detailed. This has implications for how the abstractions are used in planning and will enable empirical verification, tuning and even learning. Not too much on discovery, that should not even come up. Some stuff probably on exploration though.

Planning in the maze will involve only two or three steps at the abstract level. That is the point—that planning can be easy with the right abstractions, and the right abstractions can be expressed experientially and completely in terms of the minimal ontology.

The end. the rest of this is just notes to self.

For futures there are further possibilities. Predictive questions can be hypothetical, or subjunctive, that is, they can be conditional on particular actions being taken. The simplest case is the question about a specific observation conditional on a specific action: given a history up through an observation, what is the probability that the next observation would be o if the next action were a ?

Can questions be hypothetical on observations? I think yes.

All questions are contingent on a particular hypothetical way of behaving. They are subjunctive.

A prediction is a loss function. Minimize and know its expectation given a way of behaving. And perhaps given a way of outcoming. If certain things happen, all bets are off, just as they are if the wrong actions are taken. Events can produce divergence.

Perhaps all can be broken into next step predictions...in a sense. Recall how any environment can be described in two ways. We can give probabilities to complete sequences, and these define probabilities for all events. But this approach gets a little sticky as we introduce actions. The other way is to give probabilities to each next observation, where the probabilities depend on the experience up to that time (just as policies give probabilities for each next action, and may be non-Markov).

So the idea now is that multi-step predictions are focused anyway on the next prediction. They are focused on it and on the remaining part, just as TD methods focus on the immediate observation and the part remaining. If the immediate part is high, then less is needed from the remaining; if low, then more is needed. Perhaps any prediction of interest can be divided into an immediate part and an upcoming part. This does not necessarily mean TD methods, just a separation. Or perhaps it does mean TD methods, just not necessarily bootstrapping ones.

So, how can we structure our loss functions in these terms? I need a way of trading disappointments/confirmations between now and later. If I expect something in the exponentially discounted future, and it does not happen immediately, then I have to expect more in the exponentially discounted future from that time on. If I expect something before termination, and it does not come and I terminate a little, then I have to expect even more thereon. (and there are obvious bootstrapping possibilities.)

So what is the space of questions? What is a question? What is the space of loss functions?

The special thing about life is that it has a *now*. A centering around an instant, with a before and an after, but really a now. Now is when decisions must be made. Now is when one has the fullest representation of the current situation. AI and learning mechanisms should be oriented around the now, as TD methods are. There is a now; predictions are made and verified in an ever advancing now. Their structure is based on this. Confirmation must be done piecemeal, a time step at a time. Things are seen and new predictions are made. These must align. The verification must be step-by-step. All questions must break down this way.

Ok, to proceed. I have a formalization of general questions. Given a way of behaving, a policy, we have a probability distribution over futures. For each future, we can measure something, where here a future is not restricted to the time after now, though it could be wlog. Thus, we can talk about the expected value of that measurement. Or, more generally, there could be a loss for each future, and then we have an expected loss. The loss could have an additive structure, or not. That is, it

could be a sum over losses on each step, which is convenient for td and for learning. Otherwise general case we wait until the end and only then do an update.

Then we have a specific, now-centered form, a td form of question. There is a prediction at each time of a measurement at the next time. These are incremental questions. A prediction and a one-step loss function, where the loss is a function not just of the next observation, but of the next predictions as well.

Another general form, a complete form, is to make one-step questions about the next observation.

multi-step.

options, subjunctive

predictive state reps

TD nets - compositional

related work

conclusion

Acknowledgments

References

Sutton, R.S., Barto, A.G. (1998). *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA.