Theoretical Impediments to Machine Learning
A position paper

Judea Pearl, University of California, Los Angeles
November 2016

Abstract

Current machine learning systems operate, almost exclusively, in a purely statistical mode, which puts severe theoretical limits on their performance. We consider the feasibility of leveraging counterfactual reasoning in machine learning tasks, and to identify areas where such reasoning could lead to major breakthroughs in machine learning applications.

Scientific Background

If we examine the information that drives machine learning today, we find that it is almost entirely statistical. In other words, learning machines improve their performance by optimizing parameters over a stream of sensory inputs received from the environment. It is a slow process, analogous in many respects to the evolutionary survival-of-the-fittest process that explains how species like eagles and snakes have developed superb vision systems over millions of years. It cannot explain however the super-evolutionary process that enabled humans to build eyeglasses and telescopes over barely one thousand years. What humans possessed that other species lacked was a mental representation, a blue-print of their environment which they could manipulate at will to *imagine* alternative hypothetical environments for planning and learning. Anthropologists like N. Harari, and S. Mithen are in general agreement that the decisive ingredient that gave our homo sapiens ancestors the ability to achieve global dominion, about 40,000 years ago, was their ability to sketch and store a representation of their environment, interrogate that representation, distort it by mental acts of imagination and finally answer "What if?" kind of questions. Examples are interventional questions: "What if I act?" and retrospective or explanatory questions: "What if I had acted differently?" No learning machine in operation today can answer such questions about actions not taken before. Moreover, most learning machine today do not utilize a representation from which such questions can be answered.

We postulate that the major impediment to achieving accelerated learning speeds as well as human level performance can be overcome by removing these barriers and equipping learning machines with causal reasoning tools. This postulate would have been speculative twenty years ago, prior to the mathematization of counterfactuals. Not so today. Advances in graphical and structural models have made counterfactuals computationally manageable and thus rendered meta-statistical learning worthy of serious exploration. The next section summarizes these advances and explains how barriers to counterfactual thinking can be removed.

Level (Symbol)	Typical Activity	Typical Questions	Examples
1. Association $P(y\|x)$	Seeing	What is? How would seeing X change my belief in$Y$?	What does a symptom tell me about a disease? What does a survey tell us about the election results?
2. Intervention $P(y\|do(x), z)$	Doing	What if? What if I do X?	What if I take aspirin, will my headache be cured? What if we ban cigarettes?
3. Counterfactuals $P(y_x\|x', y')$	Imagining, Retrospection	Why? Was it X that caused Y? What if I had acted differently?	Was it the aspirin that stopped my headache? Would Kennedy be alive had Oswald not shot him? What if I had not been smoking the past 2 years?

Figure 1: The ladder of causation

The Three Layer Causal Hierarchy

An extremely useful insight unveiled by the logic of causal reasoning is the existence of a sharp classification of causal information, in terms of the kind of questions that each class is capable of answering. The classification forms a 3-level hierarchy in the sense that questions at level i $(i = 1, 2, 3)$ can only be answered if information from level j $(j \geq i)$ is available.

Figure 1 shows the 3-level hierarchy, together with the characteristic questions that can be answered at each level. The levels are titled 1. Association, 2. Intervention, and 3. Counterfactual. The names of these layers were chosen to emphasize their usage. We call the first level Association, because it invokes purely statistical relationships, defined by the naked data. For instance, observing a customer who buys toothpaste makes it more likely that he/she buys floss; such association can be inferred directly from the observed data using conditional expectation. Questions at this layer, because they require no causal information, are placed at the bottom level on the hierarchy. The second level, Intervention, ranks higher than Association because it involves not just seeing what is, but changing what we see. A typical question at this level would be: What happens if we double the price? Such questions cannot be answered from sales data alone, because they involve a change in customers behavior, in reaction to the new pricing. Customer choices under the new price structure may differ substantially from that prevailing in the past. Finally, the top level is called Counterfactuals, a term that goes back to the philosophers David Hume and John Stewart Mill, and which has been given structural semantics in the SCM framework. A typical question in the counterfactual category is "What if I were to act differently," thus necessitating retrospective reasoning.

Counterfactuals are placed at the top of the hierarchy because they subsume interventional and associational questions. If we have a model that can answer counterfactual queries, we can also answer questions about interventions and observations. For example, the interventional question, What will happen if we double the price? can be answered by asking the counterfactual question:

What would happen had the price been twice its current value? Likewise, associational questions can be answered once we can answer interventional questions; we simply ignore the action part and let observations take over. The translation does not work in the opposite direction. Interventional questions cannot be answered from purely observational information (i.e., from statistical data alone). No counterfactual question involving retrospection can be answered from purely interventional information, such as that acquired from controlled experiments; we cannot re-run an experiment on subjects who were treated with a drug and see how they behave had then not given the drug. The hierarchy is therefore directional, with the top level being the most powerful one.

Counterfactuals are the building blocks of scientific thinking as well as legal and moral reasoning. In civil court, for example, the defendant is considered to be the cause of an injury to the plaintiff if, *but for* the defendant's action, it is more likely than not that the injury would not have occurred. The computational meaning of *but for* calls for comparing the real world to an alternative world in which the defendant action did not take place.

Each layer in the hierarchy has a syntactic signature that characterizes the the sentences admitted into that layer. For example, the association layer is characterized by conditional probability sentences, e.g., $P(y|x) = p$ stating that: the probability of event $Y = y$ given that we observed event $X = x$ is equal to p. In large systems, such evidential sentences can be computed efficiently using Bayesian Networks, or any of the graphical models that support deep-learning systems.

At the interventional layer we find sentences of the type $P(y|do(x), z)$, which denotes "The probability of event $Y = y$ given that we intervene and set the value of X to x and subsequently observe event $Z = z$. Such expressions can be estimated experimentally from randomized trials or analytically using Causal Bayesian Networks (Pearl, 2000, Chapter 1).

Finally, at the counterfactual level, we have expressions of the type $P(y_x|x', y')$ which stand for "The probability of event $Y = y$ had X been x, given that we actually observed X to be x' and and Y to be y'. Such sentences can be computed only when we possess functional or Structural Equation models, or properties of such models.

This hierarchy, and the formal restrictions it entails, explains why statistics-based machine learning systems are prevented from reasoning about actions, experiments and explanations. It also suggests what external information need to be provided to, or assumed by, a learning system, and in what format, in order to circumvent those restrictions.

A Proposed Counterfactual Approach to Adaptive Decision Making

Consider the instruction: "You should have acted differently," in the context of an agent who is optimizing an action strategy. Most children learn to improve behavior by responding to such instructions, be they from parents, teachers, coaches, or reflections upon one's own experience. The information value of such instructions may encapsulate hours of trial and error learning. Yet to parse this instruction, an agent must possess the tools of counterfactual reasoning which are absent from current day learning machines. The interpretation of this instruction reads: "You have acted $X = x$, your outcome was $Y = y$, but, had you acted differently, say $X = x'$, your outcome would have been better, perhaps $Y = y'$." Formally, we can write this sentence as

$$X = x \text{ and } Y = y \implies Y_{x'} = y'$$

or, in probabilistic terms:

$$P(Y_{x'} = y' | X = x, Y = y) = \textit{high}$$

The information provided by the conditioning events: $X = x$ and $Y = y$ is extremely important, since it is this information which is specific to the agent, and carries a summary of the agent's motivation, response pattern and other idiosyncratic yet otherwise unobserved features of the agent.

Counterfactual sentences of these type have been thoroughly analyzed within the SCM framework, and we now understand fairly well the conditions under which they can be estimated from data, both experimental and observational. A simpler version of this sentence, called the Effect of Treatment on the Treated (ETT) has attained significant attention in economics and epidemiology and reads:

$$ETT = E(Y_{x'} | X = x)$$

For example, in job training context, ETT describes the effect of treatment (training program) on those chosen to enroll in the program, or, more precisely, the expected earning (Y) of those enrolled and trained, had they not been trained. Clearly, ETT is more informative measure of the effectiveness of the program than the Average Treatment Effect (ATE) which compares the earning Y of those trained to the average earning over the entire untrained population. ETT focuses on the specific constituency enrolled in the program, and this consistency may not resemble the entire population. It may consist, to take an extreme case, of those who are guaranteed high earning with or without training.

In the context of individual decision making, ETT would capture an agent saying: "I am about to act $X = x$, what if I change my mind and act $X = x'$ instead?" Clearly, this situation is ubiquitous in many decision situations, especially those where the agent is a learning mood. Again, the agent's intent, $X = x$, carries important information about the agent's specific characteristics, and should not be ignored. For example, an agent saying: "I am about to check into the hospital, should I?" is likely to have different medical urgency than one chosen at random from the population.

These considerations lead to the conclusion that, in personal decision making, the proper objective function should be ETT, not ATE. In other words, actions should be chosen so as to maximize $ETT = E(Y_{x'} | X = x)$ over all actions $X = x'$, rather than maximizing $ATE = E(Y | do(x'))$ as is done in the standard literature.

To test these ideas Bareinboim, Forney, and Pearl (2015) incorporated the ETT metric in the context of the Multi-Armed Bandit (MAB) problem, which is serving as the prototypical paradigm for active machine learning. In this context, an agent attempts to play a slot machine not knowing the expected payoff of each of the available machines in the casino. Thus, the agent must balance his need to learn which machine would deliver the highest payoff with his need to exploit the payoff information available at any given time. (Exploration vs. exploitation tradeoff). The standard metric for choice in this context is ATE, and it is estimated by randomization. In other words, the agent, at any given moment, chooses a machine x' that maximizes his average reward Y as determined by an experiment in which machines were chosen at random from those available at the casino. When the ATE metric was replaced with ETT, the agent was allowed to explore machines by any strategy whatsoever, but when it came to optimization, the criterion was ETT, not ATE. As expected, simulation results showed a marked improvement in both performance and speed of convergence (Bareinboim et al., 2015).

But how can ETT be estimated from the available data collected? Fortunately, the structural framework identifies precisely the conditions under which ETT is estimable from a combination of observational and experimental data (Shpitser and Pearl, 2009). Most of those conditions require some knowledge of the model, with the exception of one: when the action is binary (corresponding to two slot machines). Moreover, in the case of non-binary action, the MAB setting permits us to conduct a post-intention randomization, namely, the agent records his/her choice of machine, pauses, conducts a randomized experiment and, then, implements an ETT-optima action which may be different from the one intended. In this way, a database is created in which the intent and the action selected may be different. This in turn allows us to choose, at any given point, an action which maximized the expected reward conditional on the current intent. We call this strategy "intent-specific optimization," and use it to demonstrate the merit of leveraging the agent's intent as a source of useful information.

Of course, intent carries valuable information only when it reflects unobserved confounders that influenced the agent's choices in the past and that were not recorded. We conjecture that such unobserved confounders are ubiquitous in most decision situations.

However, the value of intent-base optimization goes beyond its success in the multi-bandit problem. It contains, we believe, the key by which counterfactual information can be extracted out of experiments. The key is to have agents who pause, deliberate, and then act, possibly contrary to their original intent. The ability to record the discrepancy between outcomes resulting from enacting one's intent and those resulting from acting after a deliberative pause, provides the information that renders counterfactuals estimable. It is this information that enables us to cross the barrier between layer 2 and layer 3 of the causal hierarchy. Such ability is not unique to multi bandit problems. Every child undergoes experiences where he/she pauses and thinks: Can I do better? If mental records are kept of those experiences, we have experimental semantic to counterfactual thinking in the form of regret sentences "I could have done better." The practical implications of this new semantics is worth exploring.

References

BAREINBOIM, E., FORNEY, A. and PEARL, J. (2015). Bandits with unobserved confounders: A causal approach. In *Advances in Neural Information Processing Systems 28* (C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama and R. Garnett, eds.). Curran Associates, Inc., 1342–1350. <http://papers.nips.cc/paper/5692-bandits-with-unobserved-confounders-a-causal-approach.pdf>.

PEARL, J. (2000). *Causality: Models, Reasoning, and Inference*. Cambridge University Press, New York. 2nd edition, 2009.

SHPITSER, I. and PEARL, J. (2009). Effects of treatment on the treated: Identification and generalization. In *Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence* (J. Bilmes and A. Ng, eds.). AUAI Press, Montreal, Quebec.