Solving Interpretation Problems With Etcetera Abduction

Andrew S. Gordon GORDON@ICT.USC.EDU
Institute for Creative Technologies, University of Southern California, Los Angeles, CA USA

Abstract

Among the most challenging problems in Artificial Intelligence are those that require human-like abilities to make sense of ambiguous observations, to interpret events in context given a wealth of life experiences and commonsense knowledge. In the 1990s, Jerry Hobbs and colleagues demonstrated how interpretation problems can be tackled with logical abduction, a combinatorial search for the best set of assumptions that logically entails the observations. Etcetera Abduction is a new approach to ranking assumptions by reifying the uncertainty of knowledge base axioms as etcetera literals, representing conditional and prior probabilities that can be combined through logical unification. In this invited talk, I will highlight some of the features of Etcetera Abduction that make it attractive compared to alternatives, and share my perspective on the role of logic-based reasoning given current trends in machine learning research.

1. Interpretation Problems

With the rapid pace of progress in machine-learning technology and the increasing availability of large datasets, one might understandably hope/fear that human-like Artificial Intelligence is just around the corner. Instead, I believe that recent machine-learning successes have helped to highlight quintessential human reasoning abilities that are not easily cast as classification or regression problems, and thus remain unsolvable using current inductive approaches. Here, I focus on the human ability to make sense of ambiguous observations, to interpret events in context given a wealth of life experiences and commonsense knowledge.

Interpretation problems seek good explanations for sets of ambiguous observations. Solutions to interpretation problems answer the question: *Why am I seeing these things together?* In interpretation problems, both inputs and outputs are compositional. The input is always a set of observations, where each one individually has multiple possible explanations that, when combined with the possible explanations of the others, compose novel explanations that account for the full set of observations. If we are not too rigid in our definitions of observations and explanations, we can see the process of interpretation at work throughout our everyday lives. In the list that follows, I highlight some of the cognitive tasks and intellectual pursuits that require the composition of novel explanations for new sets of observations.

Language understanding: The observed sequence of ambiguous words is explained by the meaning intended by the writer (semantics). The meaning and context of discourse is explained by effect intended by the sender (pragmatics). The meaning of a horoscope, song lyric, or sacred text verse is explained by its relevance to one's life.

Behavior interpretation: The observed actions of others are explained by the unobserved mental states of the actors (Theory of Mind).

Visual scene understanding: The perceptions of the eye are explained by the interactions of light and objects. The objects in the image are explained by the physical and social context in which the image was captured. The lines of a drawing are explained by objects intended by the artist.

Abductive sciences: The configuration of objects uncovered at an archeological site are explained by the behaviors and practices of the people who lived there. More generally, the observations of archeologists, astronomers, paleontologists, and geologists are explained by the unobserved entities and processes of their best scientific theories.

Diagnosis: The medical symptoms of a patient are explained by the mechanisms of a disease. The faults of a mechanical system are explained by the malfunction of its parts. The failings of a society are explained by the practices of its peoples and the policies of its governments.

Forensics and Justice: The evidence collected at a crime scene and analyzed in a laboratory are explained by unobserved behavior of a criminal. The guilt of a defendant in light of the evidence is explained by the means, motive, and opportunity for a defendant to commit the crime.

Threat detection: The caution one takes while driving is explained by the future events that would lead to a crash. More generally, the assertion of a threat in the current context is explained by the mechanism by which entities in the context will cause harm in the yet-unobserved future.

Planning: The assertion that one will thrive in the future is explained by the plan of action that one takes next. More generally, the occurrence of a future goal state is explained by the plan that achieves it.

Reflection: One's life experiences are explained by one's identity and destiny. The meaning of life events is explained by the coherent narrative that organizes them.

2. Interpretations are Combinations of Associations

Common to all of these interpretation problems is the *combining* of information associated with each of the given observations into a coherent explanation. The reasoning challenge is to figure out which bit of information associated with an observation combines best with that of the others. In the example of language understanding, these associations include all of the distinct senses of each word in a sentence, where the best interpretation identifies one sense for each word that participates in the representation of the meaning. In the example of behavior interpretation, each of the actions of another person might be associated with numerous possible motivations when viewed in isolation, but the best interpretation identifies one motivation for each action that participates in a coherent explanation of their mental state and intentions.

We can thus characterize the cognitive process of interpretation as a sequence of subprocesses:

1. *Collect associations for each observation*: Each observation under consideration has associations with other knowledge that *explain* the observation when considered independently, e.g., the senses of words, the intentions behind actions.

2. *Search through possible combinations*: Not knowing ahead of time which bit of associated knowledge will combine best with the others, all combinations may be considered.

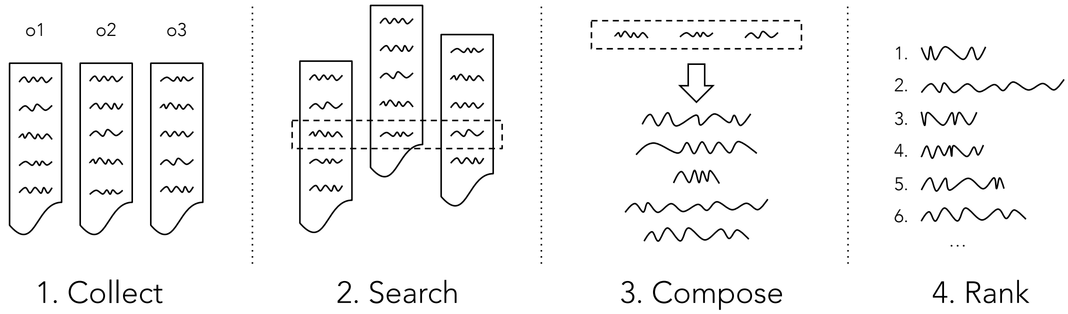

Figure 1. The Process of Interpretation

3. *Compose possible solutions from parts*: Each combination has one or more ways that the parts can be combined into a composite solution, depending on how the knowledge is represented.

4. *Rank solutions by their quality*: The relative quality of each candidate solution is judged according to some objective measure to identify the best solution(s).

Characterized in this general manner, there are a wide variety of ways that the process of interpretation could be modeled in software. At this level of abstraction, there are no commitments to particular styles of knowledge representation, search strategies, composition methods, or measures of the relative quality of solutions. However, the core *search* and *compose* steps seem particularly challenging to cast as nonlinear functions learned from training data, e.g. using contemporary machine-learning methods. Even if one day we devise a machine-learning approach to combinatorial search, we are still left with the problem of collecting the requisite training data. The correct gold-standard solutions to interpretation problems are, of course, *open to interpretation*; achieving high inter-rater agreement will be as difficult as getting jurists to agree in court cases, viewers to agree on the meaning of abstract art, or co-workers to agree on the intent of the boss's email message.

3. Etcetera Abduction

Etcetera Abduction is a technology that models the interpretation process with specific commitments to the use of formal logic as the means of representing knowledge and probability theory as the means of ranking solutions. It accepts as input a conjunction of (propositional or first-order) logical literals as observations. It produces as output a list of solutions, each of which is a conjunction of literals that, if all were assumed as true, would logically entail the input observations, ordered by their joint probability. The logical proof of the observations from the entailing assumptions answers the question, *Why am I seeing these things together?* Because the most-probable solution logically entails the observations, Etcetera Abduction also answers the question, *What is my highest estimation of the joint probability of the things I am seeing?*

The innovation of Etcetera Abduction over previous approaches to logical abduction is that solutions and their component parts are comprised entirely of *etcetera* literals, each of which reifies distinct uncertainties that must be assumed true in a logical proof of the observations. Etcetera literals enable a knowledge base of defeasible axioms to be authored as propositional or first-order definite clauses, where etcetera literals encode conditional probabilities when appearing alongside other antecedent literals, and prior probabilities when they are the sole antecedent literal in a definite clause. Given this representation and probabilistic semantics, the process of interpretation proceeds as follows:

1. *Collect associations for each observation*: Sets of entailing assumptions for each observation literal are collected by backchaining on knowledge-base axioms to a given depth until all sets of antecedents consist entirely of etcetera literals.

2. *Search through possible combinations*: Combinations are produced by exhaustively searching the cartesian product of each observation's sets of entailing assumptions.

3. *Compose possible solutions from parts*: In the case of propositional logic, a single composition is identified as the union of each component set. In the case of first-order logic, compositions are produced by considering all possible ways to unify literals in this union.

4. *Rank solutions by their quality*: Composed solutions are ranked according to their joint probability, computed as the product of the probabilities of each etcetera literal in the solution.

As a technology, the strength of Etcetera Abduction over other attempts to combine probability theory and logical reasoning is in its (relative) simplicity. Its implementation in software is straightforward and amiable to modern optimization techniques. Its grounding in probability theory allows for its straightforward integration with other probability-friendly technologies, both on the front-end as producers of the input observations and on the back-end as consumers of knowledge-based estimates of joint probability. The greatest strength of Etcetera Abduction, however, is that potential users of this technology require only basic familiarity with formal logic (entailment, definite clauses, unification, backchaining) and basic probability theory (prior, conditional, and joint probabilities). With a low barrier-to-entry for this technology, more people can spend more time focusing on the truly difficult and unsolved problem of knowledge representation, i.e., collecting and/or encoding all of the requisite knowledge to solve interpretation problems in across all of the domains described earlier. If large-scale knowledge representation efforts can be fostered and sustained, we may yet see success in automating the quintessential human capacity for interpretation.

Acknowledgements

This research was supported by the Office of Naval Research, grant N00014-16-1-2435.