INDUCTION: INFERENCE AND PROCESS

Neal J. Rothleder nealr@umich.edu

The University of Michigan
3218 EECS Building, 1301 Beal Ave., Ann Arbor, MI 48109-2122

Abstract

Despite the variety of discussions on induction, there is still no explicit definition of the concept. However, since most treatments appear to, at least implicitly, acknowledge it, it seems clear that such a definition exists. A key obstacle seems to be the confusion between what induction is, and how induction should be performed. This research proposes the idea that induction should be viewed in two distinct, but related, ways. The more general, abstract way gets to the heart of this concept definition, and treats induction purely as an inference. In this *inference view*, there is no concern for issues of efficiency, effectiveness, and other qualities of how induction might be performed. The second way to view induction is as the process which implements the inference. This paper establishes a definition of induction in the inference view. The subtleties and implications of this definition are discussed, and are also compared with key issues and published opinions of induction.

1. Introduction and Motivation

Induction has been discussed in many ways. To name a few, [Michalski, 1983], [Dietterich, 1990], [Quinlan, 1986], [Russell, 1989], [Rendell, 1986], and [Holland et al., 1986], represent a variety of different views on the subject. While each of these works (plus a long list of others) have contributed to the overall understanding of induction, they have all fallen short in one particular area: They have failed to explicitly define the concept of induction. With all the variety of discussions, it is still unclear what induction really is. Interestingly, it seems clear that each researcher appears to have (at least implicitly) an idea of what induction is. Despite this, there is no clear conceptual description of induction. The obstacle that is encountered all too often is that the ideas about what induction is get confused with the ideas of how induction should be performed.

This research proposes the idea that induction should be viewed in two distinct, but related, ways. The more general way is to see it straightforwardly as a type of inference. For this discussion, an inference is simply a mapping from one body of knowledge to another. In this *inference view*, the important characteristics are those of the two bodies of knowledge, and any necessary relationships between them. A description of induction as an inference is a description of induction at its simplest level, without the trappings of effectiveness, efficiency, and other desirable, but unnecessary, qualities. Using more formal language, induction in the inference view is a binary relation which maps one body of knowledge (the Starting Knowledge) to another (the New Knowledge), constrained by certain criteria. This New Knowledge is sometimes referred to as the “result” of induction. There is no algorithm or process in the inference view. There is nothing to “perform” or execute. The goal of this paper is to define and discuss the inference view of induction.

The second way to view induction is as a *process*. While this view also involves the starting knowledge and the new knowledge, it is most concerned with *how* induction is performed. That is, in practical applications, there must be some way to implement the mapping and restrictions which are defined in the inference view. Here is a list of some of the issues considered in the process view of induction: viewing induction as search, identifying an hypothesis space, choosing criteria to select from the possibly infinite number of choices for the new knowledge, deciding whether a particular hypothesis is plausible, deciding whether the representation language should be complete, dealing with issues of "noise", and determining the source of training examples. None of these concerns are addressed in the inference view. All of them are highly dependent on a variety of contextual issues. For example, the plausibility of an inductive result is a function of current knowledge and subjective goals which may exist. Such goals, while useful in practice, are not necessary when viewing induction as a general inference (more on the subject of goals in Section 5).

This leads to an interesting point. Many of the issues involved with induction in the "real-world" (e.g., plausibility and predictiveness of learned knowledge) are deemed irrelevant in the inference view. If we are to build machines which do induction, it seems clear that these concerns are important and that they must be taken into account. In fact, there is no claim here that these concerns are not important. Rather, the claim is that there is "something" called induction, a type of inference, which exists apart from those concerns. By analogy, it is just as there is a type of inference called deduction, which entails all types of computational problems in its implementation. Issues of which deductive operator to apply, when to apply it, and how to apply it do not change the conceptual definition of deduction. The same is true of induction. The goal here is not to deny such concerns of practicality, but to show them in their proper place. Before such concerns can be addressed, it must first be clear what induction is. That is why the inference view is defined.

Many treatments of induction seem to begin from the premise that an inference notion of induction exists, and proceed with the undertaking of understanding the process (or implementation). Often, an inference view description is given, but it is usually fairly brief, informal, and tightly bound into a description of how that induction ought to be performed. The inference description is not made explicit on its own. For example, in [Michalski, 1983], the author says, "The goal of [inductive] inference is to formulate plausible general assertions that explain the given facts and are able to predict new facts." While the basic idea here is based in the inference view, the notion that the generated assertions should be "plausible" is unimportant in the inference view. The inference view of induction is not concerned with the plausible, useful, desirable, etc. properties of the new knowledge, primarily because such concerns typically reflect subjective measures of *how* induction should be performed. As another example, consider [Russell, 1989]. While briefly describing an inference characterization of induction, Russell spends much of his time discussing how the New Knowledge will be generated, particularly in a way which will ensure it's future validity[1]. Such concerns do not fall into the inference view. So, while Michalski's and Russell's discussions both contribute to the understanding of induction, they fail to distinguish the conceptual definition (the inference view) from the process involved in implementing it.

Of course, the inherent interactions and coexistence of the two views cannot be ignored. Certainly, choices in each affect the other. A definition for the inference view necessarily imposes restrictions on the process by which such behavior is established. There reverse is true as well. It appears, however, that in most (machine learning) approaches to induction there is an, often implicit, understanding that the former reasoning is correct. That is, the inference view is the place to begin. Further, it appears that a common inference

[1]To be fair, it does not appear to be Russell's intention to describe an overall (what we call inference) definition of induction. His presentation of such a definition is made as part of his discussion on the role of knowledge in the inductive process.

framework exists among the various researchers, beyond their individual intentions. This is where this research begins. It starts with an attempt to uncover this underlying inference definition of induction. This is done by examining various discussions of induction, and attempting to extract from them the essential requirements.

2. Induction as an Inference

The goal of this section is to characterize and define induction purely as the two bodies of knowledge involved in an inference. It is the characteristics of these bodies which will determine when an inference may be called induction. It will occasionally be easier to discuss an *inductive system* (that is, a system which is capable of induction) rather than just the abstract notions. The only characteristics of such a system which will be considered or assumed are those which directly relate to induction as an inference.

2.1 Basis of the Analysis

The analysis for this section concentrates on three induction treatments which are predominantly, if unintentionally, inference treatments themselves: [Michalski, 1983], [Russell, 1989], [Muggleton, 1994]. The belief is that there is a common understanding of induction as an inference, and that it can be found underlying the various discussions of induction[2]. Since these three discussions, briefly summarized below, describe induction in what is essentially an inference view, they were used as reference points in the task of uncovering the underlying definition. This is not to say that only these sources were considered, nor was the synthesis of the inference definition restricted to ideas explicitly expressed. Rather, these treatments form the foundation for the set of inference criteria presented. Various other treatments of induction have been polled, analyzed, and used as a consistency check to ensure that the definition presented here encompasses and reflects a wide variety of opinions. Later, it will be shown how the ideas in some of these other treatments fit into the definition derived here.

<u>Michalski</u>

Most of Michalski's discussion is used to describe, in a fairly precise way, the various quantum components involved in machine induction. He does take some space, however, to describe his inductive paradigm, which centers around the goals of finding an hypothesis which implies a set of observational statements and satisfies a set of background knowledge. Informally, it seems to capture well the intuitive idea of induction. His ensuing discussion of how this paradigm can be realized (including a description of a working program) is well done. It is lacking, however, in two areas: (1) Michalski's inductive paradigm is not thoroughly discussed as a model for all types of induction. By his own admission, Michalski limits his discussion to certain kinds of induction. This, combined with the informal nature of the paradigm, leaves a variety of important questions open. (2) There is little effort made in separating the issues of the inference view of induction from the process view. Many of the key issues in Michalski's discussion deal with defining *preference criteria*. Such issues, as discussed in Section 5, are not germane to the inference view.

<u>Russell</u>

Russell is primarily concerned with how knowledge, particularly the "prior knowledge" (knowledge possessed before induction) can be effectively used to make

---

[2]It has not gone unnoticed that, essentially, induction is being used to define induction. No attempt is made to address any philosophical circularities of such an approach.

induction most efficient and effective. In his discussion, however, he puts forth a fairly straightforward theory of induction which centers around his “Type C” entailment equation:

$$\text{Prior Knowledge} + \text{New Knowledge} \models \text{Observations}$$

While Russell discusses many of what will be seen as the salient features of the inference view of induction, as mentioned above, they are often intermixed with non-relevant (process) issues, and their full impact as a definition of induction is not recognized.

<u>Muggleton</u>

Muggleton’s discussion is meant simply as an overview of Inductive Logic Programming (ILP), but it captures nicely most of the assumptions and restrictions involved in ILP. In fact, the four equations given as the only “logical requirements of an ILP system” capture some of the key issues of the inference view of induction. For this research, the basic shortcoming with such requirements come from their basis in a strictly logical setting. For this reason, in some ways they are too restrictive (e.g., the requirement that the background knowledge should not entail the observations), and in others somewhat irrelevant (e.g., see the discussion later on negative examples).

2.2 The Defining Criteria

Induction begins with a (non-empty) body of knowledge referred to as the *Starting Knowledge (StartK)*. There is an assumption (though not a strict requirement) that this StartK is correct with respect to the environment. The StartK can always be divided into two parts: a non-empty part called the *Positive Observations (Obs+)*[3] and the remainder of the StartK which will be called the *Prior Knowledge (PriorK)*. The result of induction is *New Knowledge (NewK)* with the following restrictions: The PriorK together with the NewK, should “explain” the Obs+. The NewK should not be derivable from the StartK. Neither should it be possible to derive any contradictions from the two (NewK and StartK) together. Finally, the NewK should not be empty. To state these defining characteristics more concisely, *the NewK should consistently generalize the Obs+* .

This intuitive framework and induction criteria can be made more formal. The following terms and notation are used: The symbol $\mapsto$ is used to mean a loose form of logical satisfaction[4]. The symbol $\Box$ is used to represent a contradiction. For a particular universe, the set of all statements of knowledge is denoted K. Members of this set are called *facts*. Induction (IND) is a relation of the following form:

$$\text{IND} \subseteq P(K) \times P(K) \qquad \text{where } P(K) \text{ is the power set of } K$$

That is

$$\text{IND} = \{ \ldots , < \text{StartK}_i , \text{NewK}_i > , \ldots \}$$

where

$$\text{NewK}_i \subseteq K \text{ ; StartK}_i \subseteq K \text{ ; StartK} \neq \emptyset$$

and, the following four criteria are met:

---

[3] For now, the term *positive* has no special significance. The adjective is maintained, however, since often the role of *negative* examples is discussed in induction (see Section 4).

[4] Since many induction treatments are not set within First Order Logic, it would be incorrect to use the strict, formal definition of logical satisfaction.

Satisfaction:

$$(\exists\ Obs+)\ (\ Obs+ \subseteq StartK_i\ ;\ Obs+ \neq \emptyset\ ;\ PriorK_i + NewK_i \mapsto Obs+\ )$$

$$\text{where } PriorK_i = (\ StartK_i - Obs+\ )$$

For brevity, where clear, Satisfaction is written PriorK + NewK $\mapsto$ Obs+

Non-Deducibility:

$$StartK_i \not\mapsto NewK_i$$

Consistency:

$$StartK_i + NewK_i \not\mapsto \square$$

Non-Emptiness:

$$NewK_i \neq \emptyset$$

The following subsections provide a more detailed look at the four criteria of the inference view of induction.

2.3 Satisfaction: PriorK + NewK $\mapsto$ Obs+

Induction is driven by observations. The Satisfaction criterion captures the idea that there is some subset of the starting knowledge which can be identified as the center of inductive attention. The purpose of induction is to explain this set of observations. This characteristic is clearly described in all the accounts of induction described in this research. Notice that in the formal description of the Satisfaction criterion, the Obs+ is not identified explicitly. The observations are not a separate body of knowledge which make induction a tertiary relation. Part of the *process* of induction is to determine which subset of the StartK should be the Obs+. This subject is discussed further in Section 5: Goals and Observations.

It should be noted that the entailment in Satisfaction is not always restricted to the formal description of first order logic. Many induction problems are not set within the domain of FOL. Further, some paradigms (e.g., [Salmon, 1984] or BACON.4 [Langley et al., 1987]) don't even require that the PriorK and NewK strictly entail the observations. The entailment can be "weak", meaning that the observations are only partial consequences or are only plausible given the PriorK and NewK. However, such considerations are often made for efficiency or computational reasons, and as such, hold little weight when dealing with induction as an inference.

One interesting consequence of the fact that induction is observation driven is that randomly generated knowledge, void of this purpose, is not considered induction. That is, the wisest oracle which produces true statements about the world (e.g., "All dogs bark") has not made an inductive inference unless these facts explain some set of observations. This does not mean, however, that such observations must come from the environment "outside" an inductive system. Dietterich [Dietterich , 1990] cites AM [Lenat and Brown, 1984] as an implemented system which performs induction over its own, internally generated, observations.

2.4 Non-Deducibility: StartK $\not\mapsto$ NewK

Induction should introduce new knowledge. The NewK should not be a deductive conclusion of the starting knowledge. This restriction has a few interesting consequences. First, it ensures that the NewK will cover (or explain) new, unseen observations which the StartK does not. Together with the Satisfaction criterion, this often means the NewK will *generalize* the Obs+. Second, Non-Deducibility is important when distinguishing rote learning from induction. Without this criteria, rote learning (which is taken to mean strict memorization of the observations) could be viewed as an extreme case of induction, where

NewK = Obs+. Finally, induction cannot take place if StartK $\mapsto$ K (particularly if StartK = K). That is, in the extreme case where the starting knowledge entails all there is to know, there can be nothing new to learn. Thus, Non-Deducibility cannot be satisfied.

2.5 Consistency: StartK + NewK $\not\mapsto \square$

The new knowledge that the system gains should be consistent with what it already knows. There should be no contradictions between the two. The usefulness of this restriction in a practical systems relies on a previously mentioned assumption: that the starting knowledge is correct. Induction does not permit the addition of new knowledge which contradicts what is already known. Paradigms such as Non-monotonic Reasoning [McDermott and Doyle, 1980] are designed to handle the cases where the PriorK can be withdrawn or modified. Such changes are not permitted in induction. There are a number of reasons for this such as: (1) Permitting such changes leads to a whole new set of problems centering around which changes to the StartK to make, and when are they valid. (2) Any change made to the StartK to permit adherence to Consistency may affect the validity of other induction criteria.

2.6 Non-Emptiness: NewK $\neq \emptyset$

While the fact that the NewK should not be empty seems less critical to the definition of induction than the other three criteria, it is important nonetheless. Together with Non-Deducibility, it captures the notion that induction introduces new knowledge. In Dietterich's [Dietterich, 1990] terms, there is an increase at the knowledge level. Consider the following example where the Obs+ have already been identified from the StartK:

<u>Example 1</u> PriorK = { Fido is a dog; All hamburger is beef;
All dogs eat all kinds of beef }
Obs+ = { Fido eats hamburger }

Now, if NewK = $\emptyset$ is allowed, then all the other characteristics of induction are satisfied. But, no new knowledge has been introduced. Further, as mentioned in the discussion of Non-Deducibility, the NewK should cover unseen examples which the StartK does not. Clearly, this cannot be the case if NewK = $\emptyset$. Continuing with the example, it is still possible to satisfy all four criteria:

NewK = { All dogs eat all kinds of meat; Beef is a kind of meat }

Notice that Example 1 happens to represent a special case. In this case, any NewK which satisfies Consistency will automatically satisfy Satisfaction (since it is already the case that PriorK $\mapsto$ Obs+). Thus, when the NewK is further restricted to satisfy Non-Deducibility, it can take on completely irrelevant values while still satisfying the four induction criteria. For instance, in Example 1 above, NewK = { Neal likes to swim in October } satisfies the induction criteria. In general, whenever PriorK $\mapsto$ Obs+, *any* non-empty NewK which satisfies Consistency and Non-Deducibility gives an inductive result.

3. Other Treatments of Induction

The next task is to see how the definition presented here compares with the ideas put forth in other discussions of induction. The purpose of this section is not to completely summarize and analyze these discussions, but rather to show how the four criteria reflect the underlying ideas in these treatments, and conversely, how some of the ideas presented in these treatments reflect those captured in the four inference criteria.

Quinlan's ID3

Quinlan's ID3 [Quinlan, 1986] is not an attempt to define induction or even to characterize it in any general way. It is fairly well known, however, and a fairly well established basis for many machine learning induction systems. It is informative to see then, in a fairly straightforward manner, that ID3 meets the induction criteria. Since ID3 does not incorporate domain knowledge, the StartK is simply the set of training examples. This can be divided such that the Obs+ are also the training examples and the PriorK is empty. The NewK is the decision tree which gets generated (or the rules which the tree represents). Clearly, these rules represent a theory which explains the observations[5]. Further, unless the training examples included all possible representable examples, then the Non-Deducibility criterion will be satisfied. The Non-Emptiness criterion is trivially satisfied. Finally, provided the training data is noise-free, the decision tree (NewK) should be perfectly consistent with the StartK (Obs+).

Dietterich's Knowledge Level Learning Discussion

Dietterich [Dietterich, 1990] describes induction as a subset of Non-Deductive Knowledge Level Learning (NKLL). In a system which performs NKLL, there must be an increase at the knowledge level (Non-Emptiness criterion) and this new knowledge should not be deductively derivable from the starting knowledge (Non-Deducibility criterion). It is not surprising that Dietterich makes no mention of conditions similar to the Satisfaction and Consistency criteria. His goal is not to describe a complete theory of induction. Rather, it is to differentiate different types of learning by using the knowledge level as a reference point. Dietterich provides necessary conditions on induction, but does not go so far as to claim they are sufficient as well. Interestingly, as far as can be determined from the descriptions given, all of the systems which Dietterich cites as performing induction appear to obey both the Satisfaction and Consistency criteria.

Also of interest in Dietterich's work is the way in which his *Bias Conjecture* is reflected in the Non-Deducibility criterion. Dietterich claims that if the biases could be ascribed to the background knowledge of a learning system, then there would be no NKLL (or induction), since the learning would be deductively describable at the knowledge level. Notice the parallel between Dietterich's argument and the following: if the biases could be described in such a way as to ascribe them to the StartK, then it would be the case that StartK $\mapsto$ NewK, which would violate the Non-Deducibility criterion, and the inference would not inductive. This issue is discussed in greater detail in Section 5: The Knowledge of the Process. For now, it is enough to see how Dietterich's ideas are reflected in the inference criterion.

Rendell

Rendell [Rendell, 1986] gives an account of induction which centers around the idea of class formation. He ultimately defines the problem of induction as the problem of discovering a *utility function*, u, the purpose of which is to establish class membership. This is somewhat similar to trying to meet the Satisfaction criterion. If the Obs+ are the individual objects, **x**, (where **x** is a vector of attribute values, as described by Rendell), and their class designation, then the function u(**x**), which maps sets of values to a class, is the NewK which

[5]To obtain something closer to strict logical satisfaction, the PriorK would contain all the attribute values for the each of the training instances, and the Obs+ would contain the class assignment for each instance. In this way, the Obs+ would be a logical consequence of the PriorK + NewK.

entails the observations[6] . That is, the goal of finding the utility function is a special case of trying to meet the Satisfaction criterion. Issues of Non-Deducibility and Consistency are not addressed in Rendell's discussion. Most of it is focused on an overview of means for improving the effectiveness and efficiency of induction (process view concerns).

Holland, Holyoak, Nisbett, & Thagard

Holland et al. [Holland et al., 1986] take a different approach from most others. For an inference definition, they provide littler more than "...induction, which we take to encompass all inferential processes that expand knowledge in the face of uncertainty." For them, induction is an ongoing *process* of building and maintaining models of the environment which will enable a system to achieve its goals. The emphasis of their approach is on essential procedures and mechanisms which should be used to in order to do induction properly.

Despite this predominantly process view, there is some evidence that the four induction criteria are being met, if somewhat subtly, in Holland et al.'s paradigm. For example, consider that the goal for their inductive systems is to model the environment so that the system's goals can be effectively achieved. In their paradigm, if the current situation cannot be suitably "explained" by the knowledge base, the inductive processes will be triggered to generate candidate knowledge to account for the problem. In other words, if the Obs+ are the current situation, and the PriorK is the knowledge existing before encountering the current situation, then the goal of the inductive mechanisms is to derive NewK such that the Satisfaction criterion holds.

It also seems clear that the Non-Deducibility criterion is met, although this is not mentioned explicitly. However, a distinction is made in Holland et al.'s discussion between modifying old rules and *generating new ones*. From the examples given of such generation, it appears that these are not simply deductive conclusions from previous knowledge.

4. Recurring Issues in Induction

Since the definition of induction presented here is somewhat abstract and formal, it is useful to see how it integrates with some of the issues commonly associated with discussions of induction.

Induction as Generalization

Often, induction is viewed as obtaining some generalization. Intuitively, the NewK provides a general theory which explains a set of observations. In the inference definition of induction, if the NewK is, in fact, a generalization of the Obs+, then the NewK will meet both the Satisfaction and Non-Deducibility criteria (really only leaving Consistency).

Note however, that there is nothing in the definition which restricts the NewK to be a strict generalization of the Obs+. It is possible to satisfy the four induction criteria without the NewK containing any kind of "more general" knowledge. Consider the following example:

Example 2: PriorK = { All dogs bark. }
Obs+ = { Fido barks. }
NewK = { Fido is a dog. }

All four inductive criteria are met here. However, the NewK is a not a generalization at all. It is a very specific fact. This particular type of example, where a general rule which entails the observations is part of the PriorK, is often cited as a case of *abductive* reasoning ([Fann,

[6]The previous footnote (involving satisfaction of the Obs+) is relevant here as well.

1970], [Levesque, 1989]). This leads to interesting questions (which are not addressed here) regarding the relationship of abduction to induction. For now, conclude simply that while it does not have to be so, a NewK which is a generalization of the Obs+ makes satisfying the induction criteria easier.

Uncertainty

Another claim made of induction is that it introduces uncertainty into the knowledge base, or that the truth of premises does not guarantee the truth of the conclusions [Salmon, 1984]. In the inference definition of induction, this uncertainty is related to the Non-Deducibility criteria. That is, since the NewK is not derivable from the StartK, the NewK's correctness is unknown. Notice, this does not mean there is no way to determine the NewK's correctness. For instance, the *process* by which the NewK is derived may, in fact, guarantee it's validity. Since the inference view ignores any such process however, there is some uncertainty to the NewK with respect to the existing knowledge.

Induction as Search

A common paradigm for machine learning is that the new knowledge is obtained by searching for a desired answer in a hypothesis space [Simon and Lea, 1990], [Russell, 1989]. While such a view is integral to the *process* of induction, it has little importance in this inference view.

Negative Examples

Often, in descriptions of learning by induction (particularly those in machine learning), part of the StartK is given to *negative observations* (see, for example, [Quinlan, 1986]). Such observations, given in a task where the goal is to derive concept descriptions from examples, are descriptions of examples of objects which are *not* in the goal concept. What role do such negative observations have in the four induction criteria? As an explicit set, their role is very small. The purpose of such negative observations is one of efficiency and practicality. That is, the negative examples are meant to prune the concept (hypothesis) space (see Induction as Search) and make concept derivation easier and faster. In this way, they may have a significant influence on the way in which induction is performed (i.e., they affect the process). Of course, since these negative observations are part of the StartK, the Consistency criterion requires that the NewK will not contradict them[7].

5. Other Issues in Induction

There are two other issues involved in this discussion which are not so obvious as those in Section 4. While not included in the inference view, they are closely tied to it, and thus merit some discussion.

The Knowledge of the Process

There is another assumption underlying this inference definition: the StartK is domain knowledge. More specifically, knowledge regarding the process of induction (meta-knowledge) is *not* included in the StartK. Recall that in the inference view, induction is nothing more than a mapping. Since there is no process, there cannot be any knowledge of it. But what about in actual, implemented systems? Where does the knowledge which embodies the process (including such things as *bias*) belong? It seems that the only

[7]Such consistency is all some induction paradigms require with respect to the negative examples. Many Inductive Logic Programming efforts (as summarized in [Muggleton, 1994]) have such a condition, for example.

reasonable place to assign it is to the PriorK (part of the StartK) However, then the NewK is a logical (deductive) consequence of the StartK (since the entire "program" is now a part of the StartK). Then Non-Deducibility is not satisfied, and there is no induction. The implication this has for machine learning is that, so far, computers have not achieved the ability to do induction! Since every result of a computer program is a logical consequence of its syntax and data, the Non-Deducibility criterion cannot be met. This is not simply a trivial result of the fact that computers can be seen as truth-preserving deduction engines. Indeed, there are implications beyond those of machine learning. It starts to become apparent that it is quite difficult to infer genuinely new knowledge. This leads back to where this research started. If machines cannot do induction, the question becomes, "What can do induction?" This leads to the questions "What is induction?" This is the question which must be answered first. This paper is an attempt to do so.

Goals and Observations

An important theme (discussed further in Section 2.3: Satisfaction) is that induction is driven by the need to explain some set of observations. As described earlier, these observations are a subset of the starting knowledge. The obvious concern is how the Obs+ are chosen from the larger body of StartK. In practice, the observations can come from any number of sources. The important point is that they are the facts for which a "theory" is needed. They may be literal "observations" in the form of some type of sensory input. They may be some part of a larger body of knowledge. They could even be some notion thought up in a daydream. Whatever their origin, however, in any kind of *practical application*, there are *goals* which motivate the choice of Obs+. Occasionally, these goals are obvious and explicit. For example, humans often study examples of physics phenomena in an attempt to derive the underlying theory. Often however, the goals are implicit. For example, it is often the human user who imposes his/her goals onto the selection of the observations. In either case, the observations are a reflection of the current goals. One of the reasons this relationship is quite pervasive in practical induction is that there are efficiency efficacy advantages to properly choosing the relevant Obs+. In general, in the inference view of induction, however, such a direct link between observations and goals is neither explicitly necessary nor intentional by design.

This leads to an obvious question: What about induction not based on practical, rational systems? If there are no identifiable goals (either implicit or explicit), can there still be induction? There answer is yes. As it has been defined here, an inductive inference has been made when the four criteria have been satisfied. Thus, while there is almost always a reason to choose a particular set of knowledge as the Obs+, the selection could just as well be random. Such a random choice of observations might serve no real purpose, but it could still lead to induction.

6. Conclusions and Future Work

There are two ways to view induction: as an inference, and as a process. Despite the previous confusion, it is possible to define induction solely as an inference, with reference only to the two bodies of knowledge involved, and without reference to issues of the process. The four induction criteria established here; Satisfaction, Consistency, Non-Deducibility, and Non-Emptiness; represent the necessary and sufficient conditions for defining induction as an inference. This definition synthesizes, encompasses, and reflects the common ideas in a variety of induction treatments. Through the criteria, the overall purpose and goals of induction have been identified, and various issues and paradigms of induction have been explained. Some of the key points:

• Induction can be viewed as an inference; mapping one body of knowledge to another.

• Induction is driven by an implicit goal to explain some set of observations.

• Induction should introduce new knowledge. That is, knowledge which is not a logical consequence of the starting body of domain knowledge.

• The new knowledge obtained in induction should be consistent with its starting knowledge. Induction does not permit modification of the starting knowledge.

• While generalization can make satisfying the induction criteria easier, it is not strictly required for induction.

Armed with this definition of induction as an inference, it would be useful to better establish the relationship with other types of inference and learning (e.g., deduction, abduction, circumscription, and discovery). Additionally, an obvious next step is to explore the Process View of Induction in a manner similar to that taken to establish the inference definition. There are clearly some restrictions placed on the process by the inference definition. For example, there must be some way to identify what subset of the StartK will be the Obs+. Separate from, and in addition to, these restrictions, it seems there should be some underlying characterization of how induction ought to be performed (particularly by machines). Further, does such a process really lead to induction as defined by the inference view? Continuing work is being applied in both these areas.

If the observation made in Section 5: The Knowledge of the Process, is true, that computers cannot perform induction, an interesting philosophical and psychological exploration is to examine if any “real” system can do induction, particularly humans. If not, what does this say about the nature of inductive inference?

Acknowledgments

The author would like to thank Keki Irani for helpful and interesting discussions on the topics in this paper.

References

[Dietterich, 1990] Thomas G. Dietterich. Learning at the knowledge level. In Jude W. Shavlik and Thomas G. Dietterich, editors, *Readings in Machine Learning*. Morgan Kaufman Publishers, Inc., 1990.

[Fann, 1970] K.T. Fann. *Peirce’s Theory of Abduction*. Martinus Nijhoff, 1970.

[Helft, 1989] Nicolas Helft. Induction as nonmonotonic inference. In *Proceedings KR ‘89*, Toronto, Ontario, 1989.

[Holland *et al*., 1986] John H. Holland, Keith J. Holyoak, Richard E. Nisbett, and Paul R. Thagard. *Induction: Processes of Inference, Learning, and Discovery*. MIT Press, 1986.

[Langley et al., 1987] Pat Langley, Herbert A. Simon, Gary L. Bradshaw, and Jan M. Zytkow. *Scientific Discovery, Computational Explorations of the Creative Processes*. The MIT Press, 1987.

[Lenat and Brown, 1984] D.B. Lenat and J.S. Brown. Why AM and EURISKO appear to work. *Artificial Intelligence*, 23: 269-294, 1984.

[Levesque, 1989] Hector J. Levesque. A knowledge-level account of abduction. In Proceedings, Eleventh IJCAI, pp. 1061-1067, Detroit, Michigan, August 1989.

[McDermott and Doyle, 1980] D. McDermott and J. Doyle. Non-monotonic Reasoning I. *Artificial Intelligence*, 13: 41-72, 1980.

[Michalski, 1983] Ryszard S. Michalski. A theory and methodology of inductive learning. In Ryszard S. Michalski, Jaime G. Carbonell, and Tom M. Mitchell, editors, *Machine Learning: An Artificial Intelligence Approach.* Morgan Kaufman Publishers, Inc., 1983.

[Muggleton, 1994] Stephen Muggleton. Inductive Logic Programming: derivations, successes and shortcomings. *SIGART Bulletin*, 5 (1), 5-11, 1994

[Quinlan, 1986] J. R. Quinlan. Induction of decision trees. *Machine Learning*, 1:81-106, 1986.

[Rendell, 1986] Larry Rendell. A General Framework for Induction and a Study of Selective Induction. *Machine Learning,* 1:177-226, 1986.

[Russell, 1989] Stuart J. Russell. *The Use of Knowledge in Analogy and Induction*. Morgan Kaufamn Publishers, Inc., 1989.

[Simon and Lea, 1990] Herbert A. Simon and Glenn Lea. Problem solving and rule induction: A unified view. In Jude W. Shavlik and Thomas G. Dietterich, editors, *Readings in Machine Learning*. Morgan Kaufman Publishers, Inc., 1990.

[Salmon, 1984] Wesley C. Salmon. *Logic,* chapter three. Prentice-Hall, Inc., 1984.