The Induction Problem in First Order Theories

Alvaro J. Riascos
University of los Andes and Quantil

August 18, 2014

Contenido

Introduction: Motivation

- The reliability of scientific inquiry and its methods has long been discussed and its a central topic in the philosophy of science (e.g., Hume, Kant, Popper, Goodman, etc.).
- We ask if there is a formal set up in which we can logically entail that a method of inquiry will in some sense converge to the truth given some background assumptions (possible worlds).
- We call this the induction problem.

Introduction: Motivation

- The reliability of scientific inquiry and its methods has long been discussed and its a central topic in the philosophy of science (e.g., Hume, Kant, Popper, Goodman, etc.).
- We ask if there is a formal set up in which we can logically entail that a method of inquiry will in some sense converge to the truth given some background assumptions (possible worlds).
- We call this the induction problem.

Introduction: Motivation

- The reliability of scientific inquiry and its methods has long been discussed and its a central topic in the philosophy of science (e.g., Hume, Kant, Popper, Goodman, etc.).
- We ask if there is a formal set up in which we can logically entail that a method of inquiry will in some sense converge to the truth given some background assumptions (possible worlds).
- We call this the induction problem.

Introduction: Motivation

- Recent progress in testability and rationality theories in economics motivate natural approaches to the induction problem.
- In particular, every notion of rationality (consistency of choices with finite data sets) provides an instance of the induction problem.

Introduction: Motivation

- Recent progress in testability and rationality theories in economics motivate natural approaches to the induction problem.
- In particular, every notion of rationality (consistency of choices with finite data sets) provides an instance of the induction problem.

Introduction: Motivation

- We think it is valuable, at least for a personal relief (if any), to ask how reliable are our methods of inquiry.
- More seriously, as Popper puts it:

  ...only a revival of interest in these riddles (of the world and man's knowledge of that world) can save the sciences and philosophy from an obscurantist faith in the expert's special skill and his personal knowledge and authority.

Introduction: Motivation

- We think it is valuable, at least for a personal relief (if any), to ask how reliable are our methods of inquiry.
- More seriously, as Popper puts it:

  ...only a revival of interest in these riddles (of the world and man's knowledge of that world) can save the sciences and philosophy from an obscurantist faith in the expert's special skill and his personal knowledge and authority.

Introduction: What we do...

- We introduce a general model of rationalizable data sets and pose the induction problem in the simplest interesting logic: First order logic.
- The connection between mathematical logic (computability theory) and the induction problem was introduced by H. Puntman.
- Special cases and/or similar formalizations within first order logic have already been proposed in the literature:
  - Shapiro (1982). Inductive Inference of Theories from Facts.
  - Osherson, Weinstein (1983). Formal Learning Theory.
  - Glymour (1985). Inductive Inference in the Limit.
  - Kelly (1996). The Logic of Reliable Inquiry.

Introduction: What we do...

- We introduce a general model of rationalizable data sets and pose the induction problem in the simplest interesting logic: First order logic.
- The connection between mathematical logic (computability theory) and the induction problem was introduced by H. Puntman.
- Special cases and/or similar formalizations within first order logic have already been proposed in the literature:
  - Shapiro (1982). Inductive Inference of Theories from Facts.
  - Osherson, Weinstein (1983). Formal Learning Theory.
  - Glymour (1985). Inductive Inference in the Limit.
  - Kelly (1996). The Logic of Reliable Inquiry.

Introduction: What we do...

- We introduce a general model of rationalizable data sets and pose the induction problem in the simplest interesting logic: First order logic.
- The connection between mathematical logic (computability theory) and the induction problem was introduced by H. Puntman.
- Special cases and/or similar formalizations within first order logic have already been proposed in the literature:
  - Shapiro (1982). Inductive Inference of Theories from Facts.
  - Osherson, Weinstein (1983). Formal Learning Theory.
  - Glymour (1985). Inductive Inference in the Limit.
  - Kelly (1996). The Logic of Reliable Inquiry.

Introduction: How we differentiate...

- Most of this literature focuses on the computational aspects of learning machines for inferring recursive functions from finite samples of their graphs.
- We rather focus on other central concepts of the induction problem: the data set which we allow scientists to observe (the rationalizable data sets).
- Partial observability (Glymour, Echenique, et.al) and weaker notions of rationalizable data sets (Caicedo, et.al) relevant in economics (decision theory, revealed preference, approximate consistency, aggregate matching, etc.) suggests a specific form of the induction problem in first order logic.

Introduction: How we differentiate...

- Most of this literature focuses on the computational aspects of learning machines for inferring recursive functions from finite samples of their graphs.
- We rather focus on other central concepts of the induction problem: the data set which we allow scientists to observe (the rationalizable data sets).
- Partial observability (Glymour, Echenique, et.al) and weaker notions of rationalizable data sets (Caicedo, et.al) relevant in economics (decision theory, revealed preference, approximate consistency, aggregate matching, etc.) suggests a specific form of the induction problem in first order logic.

Introduction: How we differentiate...

- Most of this literature focuses on the computational aspects of learning machines for inferring recursive functions from finite samples of their graphs.
- We rather focus on other central concepts of the induction problem: the data set which we allow scientists to observe (the rationalizable data sets).
- Partial observability (Glymour, Echenique, et.al) and weaker notions of rationalizable data sets (Caicedo, et.al) relevant in economics (decision theory, revealed preference, approximate consistency, aggregate matching, etc.) suggests a specific form of the induction problem in first order logic.

Introduction: What are we looking for...

- Many scientific theories are motivated and constructed using an axiomatic approach (decision theory is a good example).
- An axiomatic approach provides: a unifying framework, motivates generalizations, raise and link decidability and computability issues among others.

Introduction: What are we looking for...

- Many scientific theories are motivated and constructed using an axiomatic approach (decision theory is a good example).
- An axiomatic approach provides: a unifying framework, motivates generalizations, raise and link decidability and computability issues among others.

Introduction: What are we looking for...

- We want to characterize syntactically when is it that an induction problem is solvable (in some sense), given a class of proposed models (possible worlds).
- Some examples are given to illustrate the difficulties.
- Also, we show how to make progress in specific cases that suggest how would the solution to the general problem look like.
- Interestingly, the general problem has not been as widely studied as one would expect in the model theoretic literature (under a different name).

Introduction: What are we looking for...

- We want to characterize syntactically when is it that an induction problem is solvable (in some sense), given a class of proposed models (possible worlds).
- Some examples are given to illustrate the difficulties.
- Also, we show how to make progress in specific cases that suggest how would the solution to the general problem look like.
- Interestingly, the general problem has not been as widely studied as one would expect in the model theoretic literature (under a different name).

Introduction: What are we looking for...

- We want to characterize syntactically when is it that an induction problem is solvable (in some sense), given a class of proposed models (possible worlds).
- Some examples are given to illustrate the difficulties.
- Also, we show how to make progress in specific cases that suggest how would the solution to the general problem look like.
- Interestingly, the general problem has not been as widely studied as one would expect in the model theoretic literature (under a different name).

Introduction: What are we looking for...

- We want to characterize syntactically when is it that an induction problem is solvable (in some sense), given a class of proposed models (possible worlds).
- Some examples are given to illustrate the difficulties.
- Also, we show how to make progress in specific cases that suggest how would the solution to the general problem look like.
- Interestingly, the general problem has not been as widely studied as one would expect in the model theoretic literature (under a different name).

Contenido

Inquiry concerning first order theories

- An induction problem in first order theory is $(\mathcal{K}, \mathcal{H}, \mathcal{C}, \mathcal{M})$ where:

1. $\mathcal{K}$ is a class of L structures. Possible worlds or background knowledge (e.g., restrictions of the experimental set up).
2. $\mathcal{H}$ is a set of L sentences. Propositions under investigation.
3. $\mathcal{C} \times \mathcal{H} \to \{0,1\}$ is a correctness function. In first order logic we use the standard truth relation $\models$.
4. $\mathcal{M}$ is a set of *assessment* methods based on observable data (finite data sets) and the proposition under investigation.

Inquiry concerning first order theories

- An induction problem in first order theory is $(\mathcal{K}, \mathcal{H}, \mathcal{C}, \mathcal{M})$ where:
  1. $\mathcal{K}$ is a class of L structures. Possible worlds or background knowledge (e.g., restrictions of the experimental set up).
  2. $\mathcal{H}$ is a set of L sentences. Propositions under investigation.
  3. $\mathcal{C} \times \mathcal{H} \rightarrow \{0, 1\}$ is a correctness function. In first order logic we use the standard truth relation $\models$.
  4. $\mathcal{M}$ is a set of *assessment* methods based on observable data (finite data sets) and the proposition under investigation.

Inquiry concerning first order theories

- An induction problem in first order theory is $(\mathcal{K}, \mathcal{H}, \mathcal{C}, \mathcal{M})$ where:
  1. $\mathcal{K}$ is a class of L structures. Possible worlds or background knowledge (e.g., restrictions of the experimental set up).
  2. $\mathcal{H}$ is a set of L sentences. Propositions under investigation.
  3. $\mathcal{C} \times \mathcal{H} \to \{0, 1\}$ is a correctness function. In first order logic we use the standard truth relation $\models$.
  4. $\mathcal{M}$ is a set of *assessment* methods based on observable data (finite data sets) and the proposition under investigation.

Inquiry concerning first order theories

- An induction problem in first order theory is $(\mathcal{K}, \mathcal{H}, \mathcal{C}, \mathcal{M})$ where:
  1. $\mathcal{K}$ is a class of L structures. Possible worlds or background knowledge (e.g., restrictions of the experimental set up).
  2. $\mathcal{H}$ is a set of L sentences. Propositions under investigation.
  3. $\mathcal{C} \times \mathcal{H} \rightarrow \{0, 1\}$ is a correctness function. In first order logic we use the standard truth relation $\models$.
  4. $\mathcal{M}$ is a set of *assessment* methods based on observable data (finite data sets) and the proposition under investigation.

Inquiry concerning first order theories

- An induction problem in first order theory is $(\mathcal{K}, \mathcal{H}, \mathcal{C}, \mathcal{M})$ where:
  1. $\mathcal{K}$ is a class of L structures. Possible worlds or background knowledge (e.g., restrictions of the experimental set up).
  2. $\mathcal{H}$ is a set of L sentences. Propositions under investigation.
  3. $\mathcal{C} \times \mathcal{H} \rightarrow \{0, 1\}$ is a correctness function. In first order logic we use the standard truth relation $\models$.
  4. $\mathcal{M}$ is a set of *assessment* methods based on observable data (finite data sets) and the proposition under investigation.

- What is an assessment method:
  1. Let $\mathcal{D}$ be the class of all ω chains of L structures consistent (rationalizable) with $\mathcal{K}$.
  2. An assessment method is a function $\alpha : \mathcal{D} \times \mathcal{K} \times \mathbf{N} \to \{0, 1\}$ such that:

     For all $n \in \mathbf{N}$ and $\psi \in \mathcal{H}$, if $(\mathfrak{D}_i : i < \omega)$ and $(\mathfrak{D}' : i < \omega)$ are two chains consistent with $\mathcal{K}$ and $(\mathfrak{D}_i : i \leq n) = (\mathfrak{D}' : i \leq n)$ then:

$$\alpha((\mathfrak{D}_i : i < \omega), \psi, n) = \alpha((\mathfrak{D}' : i < \omega), \psi, n) \qquad (1)$$

- What is an assessment method:
  1. Let $\mathcal{D}$ be the class of all ω chains of L structures consistent (rationalizable) with $\mathcal{K}$.
  2. An assessment method is a function $\alpha : \mathcal{D} \times \mathcal{K} \times \mathbf{N} \rightarrow \{0, 1\}$ such that:

     For all $n \in \mathbf{N}$ and $\psi \in \mathcal{H}$, if $(\mathfrak{D}_i : i < \omega)$ and $(\mathfrak{D}' : i < \omega)$ are two chains consistent with $\mathcal{K}$ and $(\mathfrak{D}_i : i \leq n) = (\mathfrak{D}' : i \leq n)$ then:

$$\alpha((\mathfrak{D}_i : i < \omega), \psi, n) = \alpha((\mathfrak{D}' : i < \omega), \psi, n) \qquad (1)$$

- What is an assessment method:

1. Let $\mathcal{D}$ be the class of all ω chains of L structures consistent (rationalizable) with $\mathcal{K}$.
2. An assessment method is a function $\alpha : \mathcal{D} \times \mathcal{K} \times \mathbf{N} \to \{0, 1\}$ such that:

For all $n \in \mathbf{N}$ and $\psi \in \mathcal{H}$, if $(\mathfrak{D}_i : i < \omega)$ and $(\mathfrak{D}' : i < \omega)$ are two chains consistent with $\mathcal{K}$ and $(\mathfrak{D}_i : i \leq n) = (\mathfrak{D}' : i \leq n)$ then:

$$\alpha((\mathfrak{D}_i : i < \omega), \psi, n) = \alpha((\mathfrak{D}' : i < \omega), \psi, n) \qquad (1)$$

Inquiry concerning first order theories

- What is an assessment method:
  1. Let $\mathcal{D}$ be the class of all ω chains of L structures consistent (rationalizable) with $\mathcal{K}$.
  2. An assessment method is a function $\alpha : \mathcal{D} \times \mathcal{K} \times \mathbf{N} \to \{0, 1\}$ such that:

     For all $n \in \mathbf{N}$ and $\psi \in \mathcal{H}$, if $(\mathfrak{D}_i : i < \omega)$ and $(\mathfrak{D}' : i < \omega)$ are two chains consistent with $\mathcal{K}$ and $(\mathfrak{D}_i : i \leq n) = (\mathfrak{D}' : i \leq n)$ then:

$$\alpha((\mathfrak{D}_i : i < \omega), \psi, n) = \alpha((\mathfrak{D}' : i < \omega), \psi, n) \tag{1}$$

Inquiry concerning first order theories

- What is an assessment method:
    1. Let $\mathcal{D}$ be the class of all ω chains of L structures consistent (rationalizable) with $\mathcal{K}$.
    2. An assessment method is a function $\alpha : \mathcal{D} \times \mathcal{K} \times \mathbf{N} \to \{0, 1\}$ such that:

       For all $n \in \mathbf{N}$ and $\psi \in \mathcal{H}$, if $(\mathfrak{D}_i : i < \omega)$ and $(\mathfrak{D}' : i < \omega)$ are two chains consistent with $\mathcal{K}$ and $(\mathfrak{D}_i : i \leq n) = (\mathfrak{D}' : i \leq n)$ then:

$$\alpha((\mathfrak{D}_i : i < \omega), \psi, n) = \alpha((\mathfrak{D}' : i < \omega), \psi, n) \qquad (1)$$

Definition (Forms of Solvability)

- α verifies $\mathcal{H}$ (in the limit) given $\mathcal{K}$ on $\mathcal{D}$ data if for all hypothesis $\psi \in \mathcal{H}$ and for all $\mathfrak{M} \in \mathcal{K}$ and $(\mathfrak{D}_i : i < \omega) \in \mathcal{D}$ consistent with $\mathcal{K}$:

$$\mathfrak{M} \models \psi \leftrightarrow \exists n \in \mathbf{N} \text{ such that } \forall m \geq n, \alpha(\mathfrak{M}, (\mathfrak{D}_i : i < \omega), m) = 1 \tag{2}$$

- We define similarly α falsifies $\mathcal{H}$ (in the limit) given $\mathcal{K}$ on $\mathcal{D}$.
- We say that α decides $\mathcal{H}$ (in the limit) given $\mathcal{K}$ from $\mathcal{D}$ if it both verifies and falsifies $\mathcal{H}$.

Definition (Forms of Solvability)

- α verifies $\mathcal{H}$ (in the limit) given $\mathcal{K}$ on $\mathcal{D}$ data if for all hypothesis $\psi \in \mathcal{H}$ and for all $\mathfrak{M} \in \mathcal{K}$ and $(\mathfrak{D}_i : i < \omega) \in \mathcal{D}$ consistent with $\mathcal{K}$:

$$\mathfrak{M} \models \psi \leftrightarrow \exists n \in \mathbf{N} \text{ such that } \forall m \geq n, \alpha(\mathfrak{M}, (\mathfrak{D}_i : i < \omega), m) = 1 \tag{2}$$

- We define similarly α falsifies $\mathcal{H}$ (in the limit) given $\mathcal{K}$ on $\mathcal{D}$.
- We say that α decides $\mathcal{H}$ (in the limit) given $\mathcal{K}$ from $\mathcal{D}$ if it both verifies and falsifies $\mathcal{H}$.

Definition (Forms of Solvability)

- α verifies $\mathcal{H}$ (in the limit) given $\mathcal{K}$ on $\mathcal{D}$ data if for all hypothesis $\psi \in \mathcal{H}$ and for all $\mathfrak{M} \in \mathcal{K}$ and $(\mathfrak{D}_i : i < \omega) \in \mathcal{D}$ consistent with $\mathcal{K}$:

$$\mathfrak{M} \models \psi \leftrightarrow \exists n \in \mathbf{N} \text{ such that } \forall m \geq n, \alpha(\mathfrak{M}, (\mathfrak{D}_i : i < \omega), m) = 1 \tag{2}$$

- We define similarly α falsifies $\mathcal{H}$ (in the limit) given $\mathcal{K}$ on $\mathcal{D}$.
- We say that α decides $\mathcal{H}$ (in the limit) given $\mathcal{K}$ from $\mathcal{D}$ if it both verifies and falsifies $\mathcal{H}$.

Definition (Forms of Solvability)

- α verifies $\mathcal{H}$ (in the limit) given $\mathcal{K}$ on $\mathcal{D}$ data if for all hypothesis $\psi \in \mathcal{H}$ and for all $\mathfrak{M} \in \mathcal{K}$ and $(\mathfrak{D}_i : i < \omega) \in \mathcal{D}$ consistent with $\mathcal{K}$:

$$\mathfrak{M} \models \psi \leftrightarrow \exists n \in \mathbf{N} \text{ such that } \forall m \geq n, \alpha(\mathfrak{M}, (\mathfrak{D}_i : i < \omega), m) = 1 \tag{2}$$

- We define similarly α falsifies $\mathcal{H}$ (in the limit) given $\mathcal{K}$ on $\mathcal{D}$.
- We say that α decides $\mathcal{H}$ (in the limit) given $\mathcal{K}$ from $\mathcal{D}$ if it both verifies and falsifies $\mathcal{H}$.

Inquiry concerning first order theories: summing up

Given an induction problem $(\mathcal{K}, \mathcal{H}, \mathcal{C}, \mathcal{M})$

- $\mathcal{K}$ expresses how *reliable* are the assessment methods over which the assessment methods succeed.
- $\mathcal{H}$ the *range* (of applicability).
- $\mathcal{D}$ the observable data.
- Other notions of correctness and success may be of interest.

Inquiry concerning first order theories: summing up

Given an induction problem $(\mathcal{K}, \mathcal{H}, \mathcal{C}, \mathcal{M})$

- $\mathcal{K}$ expresses how *reliable* are the assessment methods over which the assessment methods succeed.
- $\mathcal{H}$ the *range* (of applicability).
- $\mathcal{D}$ the observable data.
- Other notions of correctness and success may be of interest.

Contenido

Rationalizable data sets: Data

Definition (Data Sets)

Let L' be a language with a finite number of constants and relation symbols such that $L' \subseteq L$. An L'-data set $\mathfrak{D}$ is a finite L'-structure. $\mathfrak{D} = (D, (R^{\mathfrak{D}})_{R \in L'}, (c^{\mathfrak{D}})_{c \in L'})$.

Rationalizable data sets: Consistency

Definition (Consistency of Data Sets)

A data set $\mathfrak{D}$ is consistent with an L-structure $\mathfrak{M}$ if there is an homomorphism of $\mathfrak{D}$ into $\mathfrak{M}$. We denote this by $\mathfrak{D} \to \mathfrak{M}$.

- Chambers, et.al (2013) require homomorphism to be 1-1.
- Simon et.al (1973) require isomorphic embedding.

Rationalizable data sets: Consistency

Definition (Consistency of Data Sets)

A data set $\mathfrak{D}$ is consistent with an L-structure $\mathfrak{M}$ if there is an homomorphism of $\mathfrak{D}$ into $\mathfrak{M}$. We denote this by $\mathfrak{D} \to \mathfrak{M}$.

- Chambers, et.al (2013) require homomorphism to be 1-1.
- Simon et.al (1973) require isomorphic embedding.

Rationalizable data sets: Consistency

Definition (Consistency of Data Sets)

A data set $\mathfrak{D}$ is consistent with an L-structure $\mathfrak{M}$ if there is an homomorphism of $\mathfrak{D}$ into $\mathfrak{M}$. We denote this by $\mathfrak{D} \to \mathfrak{M}$.

- Chambers, et.al (2013) require homomorphism to be 1-1.
- Simon et.al (1973) require isomorphic embedding.

Contenido

Examples

Example (The theory of simple linear orders with endpoints)

Consider the following theory in the language $L = (0, 1, <)$.

- $SOE = \{<$ is irreflexive, transitive, complete and 0 and 1 are the smallest and biggest elements of the universe$\}$.
- $DO = \{\forall x, y, (x < y) \to \exists z(x < z < y)\}$
- Then DO is not verifiable nor refutable in the limit given SOE.
- In Kelly's (1996) formalization DO is refutable given SOE

Examples

Example (The theory of simple linear orders with endpoints - cont)

Consider the following theory in the language $L = (0, 1, \leq)$.

- Let $\mathfrak{M} = ([0, 1], 0, 1, \leq), \mathfrak{N} = (\{0, ..., \frac{1}{4}, \frac{1}{2}, 1\}, 0, 1, \leq)$. Both are models of SOE.
- Take any data set consistent with $\mathfrak{M}, \mathfrak{N}$.

Example (Complete universal theories)

Let L be a language with countably many unary function symbols. The theory that states each function is $1-1$, has no finite loops and have disjoint ranges is complete and universal.

Examples

Example (Revealed preference theory)

$L = (\preccurlyeq, \prec)$.

- Rational choice (weak order) are the class of structures that are models of:
    1. $\forall x, \forall y (x \preccurlyeq y \vee y \preccurlyeq x)$
    2. Transitivity
    3. Consistency or characterization of $\prec$ in terms of $\preccurlyeq$

Examples

Example (Revealed preference theory - Cont)

Many sentences, even universal sentences such as $\preccurlyeq$ is monotonic is not verifiable or refutable in the limit given rational choice theory.

Contenido

Conclusions

- Very preliminary work aiming to formalize the problem of induction that captures recent ideas of rationality in economic theory.
- If we stick to the conceptual framework of the formal learning literature, results look very negative for the solvability of the induction problem.