

The Model Inference System

Ehud Y. Shapiro¹

Department of Computer Science
Yale University
New Haven, CT 06520

The Model Inference System is an interactive Prolog program [1] that infers theories from facts. It implements the general model inference algorithm, described in [2], specialized to infer theories in Horn form from atomic facts.

The abstract setting in which the system operates, and from which it derives its name, is the following: the system is initially given some first order language L . There is some unknown model M for L , and an oracle (user) that can supply facts and answer queries on the truth or falsity of ground atoms in M . The goal of the system is to find a finite set of true Horn clauses, which imply all true ground atoms and no false ones.

The user inputs the predicate to be inferred and the terms (predicate symbols, function symbols and constants) to be used in the intended axiomatization of this predicate. After receiving the initial input, the system starts prompting the user for new facts. At each stage, the system maintains a set of hypotheses that are sufficient to derive all true atoms and none of the false atoms supplied so far.

When an inconsistency is discovered between the hypotheses and the facts, the contradiction backtracing algorithm is invoked. In order to resolve this contradiction, the algorithm may query for the truth of atoms not supplied by the user yet. When the hypotheses can not derive some atom known to be true, the system adds to the conjecture refinements of previously refuted hypotheses.

The user sees the axiomatization as it is developed by the system. When he thinks the system has converged to a correct and sufficient axiomatization of the predicate, he can terminate the inference process. It is his responsibility to terminate the system, since it is an inherent property of an inductive inference algorithm that it can never determine whether it has converged.

Below are examples of the performance of the system. Further examples can be found in [2]. The system runs compiled using the in-core compiler of the Edinburgh Prolog-10 version 3. The computer is a DECsystem-2060, running Prolog-10 under the compatibility package PA1050. In the following, the running time in CPU seconds measured the time the system took to converge to a correct and sufficient axiomatization of the intended model. *Facts* are the number of facts read therein, and *hypotheses generated* is the number of hypotheses generated by the refinement operator during the inference. The numbers following an axiom are its sequential number as it was generated, and the number of refinement operations needed to generate it.

Exponentiation: $\text{exp}(X, Y, Z) \leftarrow X \text{ to the } Y \text{ is } Z$.

28 CPU seconds, 18 facts, 252 hypotheses generated.

$(\text{exp}(X, 0, s(0)) \leftarrow \text{true})$, 78, 2

$(\text{exp}(X, s(Y), Z) \leftarrow \text{exp}(X, Y, W), \text{times}(W, X, Z))$, 129, 2

$(\text{exp}(0, s(X), 0) \leftarrow \text{true})$, 148, 2

Member: $\text{member}(X, Y) \leftarrow X \text{ is a member of the list } Y$.

2 CPU seconds, 23 facts, 13 hypotheses generated.

$(\text{member}(X, [Y|Z]) \leftarrow \text{member}(X, Z))$, 6, 3

$(\text{member}(X, [X|Y]) \leftarrow \text{true})$, 7, 3

Subset: $\text{subset}(X, Y) \leftarrow X \text{ is a subset of } Y$.

31 CPU seconds, 82 facts, 152 hypotheses generated.

$(\text{subset}([], X) \leftarrow \text{true})$, 1, 1

$(\text{subset}(X, [Y|X]) \leftarrow \text{true})$, 32, 3

$(\text{subset}(X, [Y|Z]) \leftarrow \text{subset}(X, Z))$, 34, 3

$(\text{subset}([X], [X|Y]) \leftarrow \text{true})$, 97, 5

$(\text{subset}([X|Y], [X|Z]) \leftarrow \text{subset}(Y, Z))$, 103, 5

$(\text{subset}([X|Y], Z) \leftarrow \text{subset}(Y, Z), \text{member}(X, Z))$, 113, 3

Interchange: $\text{int}(X, Y) \leftarrow Y \text{ is a list of pairs of } X$, where the order in each pair is interchanged.

19 CPU seconds, 21 facts, 110 hypotheses generated.

$(\text{int}([], []) \leftarrow \text{true})$, 6, 1

$(\text{int}([X|Y], [Z], [[Y|X]|W]) \leftarrow \text{int}(Z, W), \text{true})$, 104, 9

Tree isomorphism: $\text{isotree}(X, Y) \leftarrow X \text{ is isomorphic to } Y$.

20 CPU seconds, 110 facts, 117 hypotheses generated.

$(\text{isotree}(X, X) \leftarrow \text{true})$, 5, 1

$(\text{isotree}(t(X, Y), t(Y, X)) \leftarrow \text{true})$, 35, 5

$(\text{isotree}(t(X, Y), t(Y, Z)) \leftarrow \text{isotree}(X, Z))$, 38, 5

$(\text{isotree}(t(X, Y), t(Z, W)) \leftarrow \text{isotree}(Y, Z), \text{isotree}(X, W))$, 41,

$(\text{isotree}(t(X, Y), t(Z, X)) \leftarrow \text{isotree}(Y, Z))$, 44, 5

$(\text{isotree}(t(X, Y), t(Z, Y)) \leftarrow \text{isotree}(X, Z))$, 103, 5

$(\text{isotree}(t(X, Y), t(Z, W)) \leftarrow \text{isotree}(Y, W), \text{isotree}(X, Z))$, 108

References

- [1] L. Pereira, F. Pereira and D. Warren.
User's Guide to DECsystem-10 PROLOG.
Technical Report 03 13/5570, Laboratorio Nacional De
Engenharia Civil, Lisbon, September, 1978.
Provisional version.
- [2] Ehud Y. Shapiro.
An Algorithm that Infers Theories from Facts.
IJCAI-81, this volume.

¹This work was supported in part by the National Science Foundation, grant No. MCS8002447.