Why

We want to talk about learning associations between objects in time or space.

Definition

Let A and B be sets. An *inductor* is a function mapping a dataset of paired records in $A \times B$ to a function from A to B.

We call the elements of A the *inputs* and the elements of B the *outputs*. A *predictor* is a function from the inputs to the outputs and the result of an input under a predictor is a *prediction*. Using this language, an inductor maps datasets to predictors. A predictor maps precepts to postcepts.

Notation

Let D be a dataset of size n in $A \times B$. Let $g : A \to B$, a predictor, which makes prediction $g(a)$ on precept $a \in A$. Let $f : (A \times B)^n \to (A \to B)$, an inductor. Then $f(D)$ is the predictor which the inductor associates with dataset D.

Other terminology

Many authorities call the inputs the *independent variables*, *explanatory variables*, *precepts*, *covariates*, *patterns* or *observations*. Similarly, some call the outputs the *dependent variables*, *postcepts*, *targets*, *outcomes* or *observational outcomes*.

Some call a predictor an *input-output* mapping, a *point*

predictor,[1], *prediction rule*, *hypothesis*, *concept*, or *classifier*. Some authors refer to a prediction as a *guess*.

Other authors refer to the set of inputs A as the *domain set* and elements of A as *instances*. Likewise, some authors refer to the argument of an inductor as the *training data*, *training set*,[2] or *training examples*.

Learning algorithms

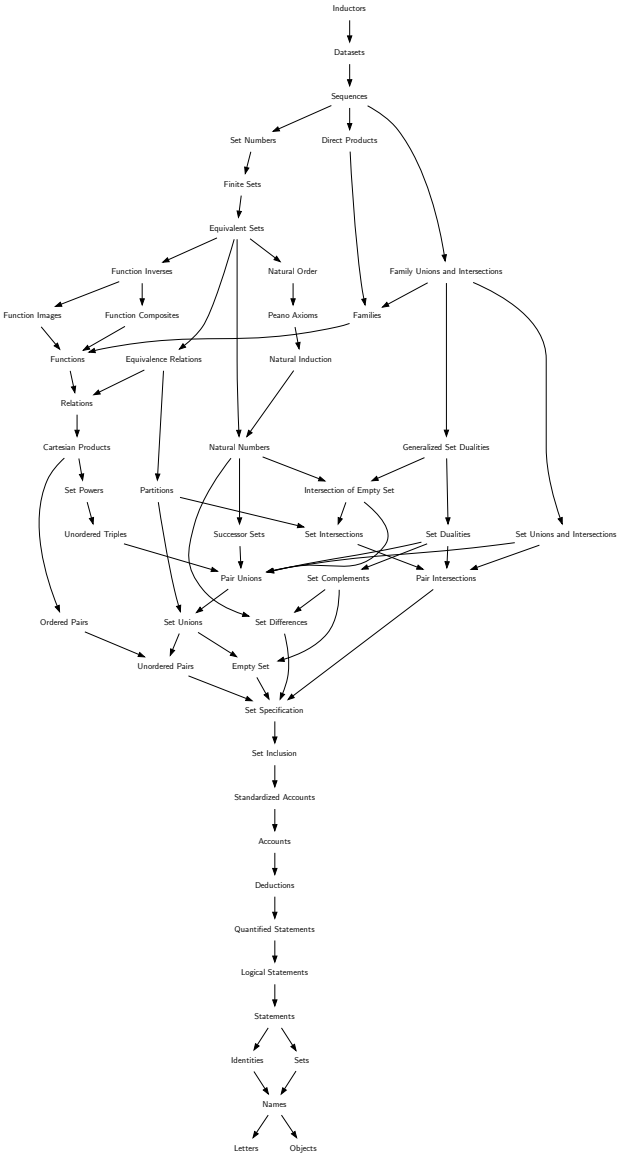
We use a predictor to make guesses on precepts which do not appear in the dataset that was used to construct the predictor. We refer to the task of proposing a predictor for a particular dataset a *learning problem*.

It is common in this context to refer to an inductor as a *learning algorithm* and call the task or problem of constructing a predictor from a dataset *supervised learning*. By supervision, we mean to indicate that we have the outputs corresponding to the inputs.

In line with this usage, the outputs are often called *labels* and the labels are said "to provide supervision." In this context, the dataset used to construct the predictor (i.e., the argument to the inductor) is called the *training dataset*.

[1]Future editions may remove this. The intuition for the word point is from the real numbers, which is not a prerequisite sheet.

[2]The word "set" is rough, since we are speaking of a sequence.


Inductors
Datasets
Sequences
Set Numbers
Direct Products
Finite Sets
Equivalent Sets
Function Inverses
Natural Order
Family Unions and Intersections
Function Images
Function Composites
Peano Axioms
Families
Functions
Equivalence Relations
Natural Induction
Relations
Cartesian Products
Natural Numbers
Generalized Set Dualities
Set Powers
Partitions
Intersection of Empty Set
Unordered Triples
Successor Sets
Set Intersections
Set Dualities
Set Unions and Intersections
Pair Unions
Set Complements
Pair Intersections
Ordered Pairs
Set Unions
Set Differences
Unordered Pairs
Empty Set
Set Specification
Set Inclusion
Standardized Accounts
Accounts
Deductions
Quantified Statements
Logical Statements
Statements
Identities
Sets
Names
Letters
Objects