Introduction

“In which we examine the problem of learning when you know something already”

Knowledge Representation + Learning
Contents

- Logical Formulation of Learning
  - Current-best-hypothesis search
  - Least-commitment search

- Knowledge in Learning
  - Explanation-based Learning (EBL)
  - Relevance-based Learning (RBL)
  - Knowledge-based Inductive Learning (KBIL)
    - Inductive Logic Programming (ILP)
Logical Formulation

- Specialize pure inductive learning to use FOL
- Aim of inductive learning in the logical setting
  - To find an equivalent logical expression for the goal predicate that can be used to classify examples correctly

- For each example
  - Description of the object: $D_i(X_i)$
    - $D_i$ can be any logical expression taking a single argument, e.g. Hungry($x$)
  - Classification of the object: $Q(X_i)$

- Training set is the conjunction of all description and classification sentences
Logical Formulation

- Each hypothesis proposes a candidate definition of the goal predicate
  - Extension of the predicate: set of examples predicted to be examples of the goal predicate
  - Two hypotheses are logically equivalent if they have the same extension
- Learning algorithm searches through the hypothesis space \( H \) to find the correct one
Logical Formulation

- As examples are considered, hypotheses that are not consistent with them can be ruled out
  - Inconsistent if the example is a false negative or false positive of the hypothesis

- Inductive learning in a logical setting:
  - Process of gradually eliminating hypotheses that are inconsistent with the examples
Search Algorithms: CBH

- Current-best-hypothesis search
  - A single hypothesis is kept and adjusted to maintain consistency as new examples come

- Adjustments
  - Generalization: extension of hypothesis is increased to include a false negative
    - By dropping conditions for a weaker definition
  - Specialization: extension of hypothesis is decreased to exclude a false positive
    - By adding conditions or removing disjuncts
Search Algorithms: CBH

Figure 19.1
- (a) A consistent hypothesis
- (b) A false negative
- (c) Hypothesis is generalized
- (d) A false positive
- (e) Hypothesis is specialized
Search Algorithms: CBH

- Algorithm is defined nondeterministically
  - A number of adjustments can be made at a point
- If an adjustment is made, examples need to be checked for consistency
- Algorithm will backtrack from irrecoverable situations

Problems
- Checking examples after each adjustment is expensive
- Search may require a lot of backtracking
  - Algorithm is forced to choose a particular hypothesis as a best guess, possibly without enough information
Search Algorithms: LCS

- Least-commitment search
  - Does not make arbitrary choices in selecting hypotheses
- Version space learning algorithm
  - Keep all the hypotheses that are consistent with the examples in a version space
  - New examples with eliminate hypotheses or have no effect on the version space
  - Incremental: will not have to check old examples
Search Algorithms: LCS

- Large hypothesis space is represented in an interval representation, like [1,2]
  - Boundary sets:
    - G-set (most general set)
    - S-set (most specific set)
  - Everything in between is guaranteed to be consistent with the examples
Search Algorithms: LCS

- As G- and S-sets are updated, hypotheses are eliminated until
  - Only one hypothesis is left
  - Version space collapses (G or S becomes empty)
    - No consistent hypothesis for this set
  - Run out of examples for exact classification
    - Version space represents a disjunction of hypotheses
Knowledge in Learning

• To understand the role of prior knowledge
  • Consider the logical relationship between hypotheses, example descriptions and classifications

• Hypothesis that explains the observations must satisfy an entailment constraint
  • $Hypothesis \land Descriptions \not\models Classifications$

• Solve the constraint by searching for the hypothesis in the hypothesis space H
Knowledge in Learning

- Task: to design agents that already knows something, and are trying to learn more
- Agents must have a learning process to gain the background knowledge in the first place
  - Learning taken place afterwards define the agent’s incremental/cumulative development
- Agents can start off like normal agents
  - Gain initial knowledge through inductive learning
  - After, uses background knowledge to learn more and more effectively
The use of prior knowledge allows for much faster learning than from pure induction learning!

Several generalization processes that use prior knowledge…
Explanation-based Learning

- Explanation-based Learning (EBL)
  - Method for extracting rules from individual observations through an explanation

- Explanation
  - Stick holds the food over the fire while keeping hands safe

- Generalization
  - Any long, rigid, sharp object can be used to toast food over the fire
  - General rule follows logically from the background knowledge of the cavemen’s usual cooking process

“Ug, Look What Zog Do.”
Gary Larson
Explanation-based Learning

\[ \text{Hypothesis} \land \text{Descriptions} \not\equiv \text{Classifications} \]

\[ \text{Background} \not\equiv \text{Hypothesis} \]

- Background knowledge is required to explain the hypothesis, which explains observations
- Knowledge is derived from what is known
  - Agent does not actually learn anything “new”
- EBL is seen as a method for converting theories into useful knowledge
  - Generalization can be used in other situations
  - Becomes obvious step to solve complex problems
Explanation-based Learning

- Extracting general rules from examples
  - Construct a proof using the given example and background knowledge, showing that the goal predicate applies to the example
    - Logical proof, reasoning/problem-solving process with well defined steps
  - In parallel, construct a generalized proof tree for the variabilized goal using the same inference steps as in the original proof
    - As the proof proceeds, some of the variables are instantiated
Explanation-based Learning

- Construct a new rule for the goal predicate
  - Left-hand side consists of the leaves of the proof tree
  - Right-hand side is the variabilized goal (after applying necessary bindings)
- Drop any conditions that are true regardless of the values of the variables in the goal
Explanation-based Learning

- Within any given generalized proof tree, more than one rule can be found
  - Rule can be extracted from any partial subtree
  - Which one to choose? Question of efficiency

- Analysis of efficiency in EBL
  - Adding many rules can slow down reasoning process
  - Derived rules must offer significant speed increase by avoiding dead ends (shortens proof)
  - Derived rules should be as general as possible to apply to the largest set of cases possible
Relevance-based Learning

- Relevance-based Learning (RBL)
  - Using only the relevant background knowledge to make generalizations
- Tourist in Brazil meets a native Brazilian
  - He speaks Portuguese!
    - Tourist concludes that all Brazilians speak Portuguese
  - His name is Fernando!
    - Tourist does not conclude that all Brazilians are called Fernando
Relevance-based Learning

- Relevant knowledge?
  - That all citizens of a country are likely to speak the same language
  - But not all citizens have the same name
- Background knowledge is concerned with the relevance of a set of features to the goal
  - Expressed as determinations or functional dependencies
    - \( \text{Nationality}(x, n) \supset \text{Language}(x, l) \)

\[
\text{Hypothesis} \land \text{Descriptions} \not\equiv \text{Classifications} \\
\text{Background} \land \text{Descriptions} \land \text{Classifications} \not\equiv \text{Hypothesis}
\]
Relevance-based Learning

- Learning algorithm for determinations
  - As all prior knowledge must be learned first
- Find the minimal consistent determination
  - The simplest determination consistent with all the observations
  - Many possible algorithms available
    - Search through the space of determinations, checking them with an increasing number of predicates to find a consistent determination
Relevance-based Learning

- If we combine RBL with a learning algorithm
  - Relevance-based Learning (RBL)
    + Decision Tree Learning Algorithm (DTL)
  - RBDTL: identifies a minimal set of relevant attributes, then passes the set to DTL
Knowledge-based Inductive Learning

- Knowledge-based Inductive Learning (KBIL)
  - Background knowledge and new hypothesis combine to explain the examples

- Medical student
  - Doctor diagnoses a patient with an infection, and tells the patient to take a particular antibiotic
    - Medical student infers that the antibiotic is effective for that particular type of infection
Knowledge-based Inductive Learning

- Knowledge?
  - Medical student needs to know enough to
    - Infer the infection from the symptoms
    - That this antibiotic is effective against it

\[ \text{Background} \land \text{Hypothesis} \land \text{Descriptions} \equiv \text{Classifications} \]

- KBIL has been studied mainly in the field of inductive logic programming (ILP)
  - Combines inductive methods with first-order representations
Inductive Logic Programming

- Prior knowledge reduces the complexity of learning in ILP systems
  - Hypothesis space size is reduced to include only theories consistent with background knowledge
  - Size of hypothesis required is reduced because background knowledge is available to help new rules explain the observations
    - Smaller hypotheses are easier to find!
Inductive Logic Programming

- ILP algorithms are constructive induction algorithms
  - Able to create new predicates to facilitate the expression of explanatory hypotheses
- Express Grandparent
  - Empty background
    - Hypotheses are long and complicated

\[ \text{Grandparent}(x, y) \iff \]
\[
\exists z \ (\text{Mother}(x, z) \land \text{Mother}(z, y)) \lor \\
\exists z \ (\text{Mother}(x, z) \land \text{Father}(z, y)) \lor \\
\exists z \ (\text{Father}(x, z) \land \text{Mother}(z, y)) \lor \\
\exists z \ (\text{Father}(x, z) \land \text{Father}(z, y))
\]
Inductive Logic Programming

- By creating a new predicate, the definition of Grandparent can be reduced

  \[ \text{Parent}(x, y) \iff [\text{Mother}(x, y) \lor \text{Father}(x, y)] \]
  \[ \text{Grandparent}(x, y) \iff [\exists z \, \text{Parent}(x, z) \land \text{Parent}(z, y)] \]

- Background knowledge can reduce the size of hypotheses required to explain the observations
Two main approaches to ILP

Top-down inductive learning method

- Start with a general rule, and gradually specialize it until it fits the data
- Algorithm repeatedly constructs a clause, one literal at a time, until it agrees with the subset of positive examples and none of the negative examples
  - Literals can be predicates, equality/inequality literals, or arithmetic comparisons
- Positive examples are removed from the set, and process continues until none remain
Inductive Logic Programming

- Type restrictions reduce the search space
  - For domains that included more than one type of information
- Application of Ockham’s razor
  - If the clause becomes longer than the total length of the positive examples that it explains, it is not considered a potential hypothesis
Inductive Logic Programming

- Inductive learning with inverse deduction
  - Inverse resolution
    - If the example classifications follow from the background, hypothesis and descriptions…
    - Then one must be able to prove this by resolution
  - Process
    - Take a resolvent $C$ to produce clauses $C_1$ and $C_2$ or
    - Take resolvent $C$ and clause $C_1$ to produce $C_2$
Inductive Logic Programming

- Inverse resolution involves a search for the clauses
  - Each step is nondeterministic because there are many/an infinite number of clauses that can be used
  - Clauses can be chosen from background knowledge, example descriptions, negated classifications, or hypothesized clauses already in the resolution
Inductive Logic Programming

- Ways to control the branching factor
  - Eliminate redundant choices
    - E.g. generate only the most specific hypotheses possible, require all clauses to be consistent with each other
  - Restrict proof strategy
  - Restrict representation language
    - E.g. do not use Horn clauses or function symbols
  - Inference can be done with model checking instead of theorem proving
    - E.g. check for consistency instead
  - Inference can be done with ground propositional clauses rather than FOL
    - Can be more efficient for some problems
Summary

- Using prior knowledge leads to cumulative learning
  - Agents improve their ability to learn as they acquire more knowledge

- EBL
  - Extract general rules from single examples by explaining them, and generalizing the explanation
  - Deductive method to turn knowledge into useful expertise

- RBL
  - Determinations identify relevant attributes to reduce the hypothesis space and speed up learning
Summary

- **KBIL**
  - Finds inductive hypotheses that explain sets of observations with background knowledge

- **ILP**
  - Has techniques that perform KBIL on knowledge expressed in FOL
  - Learns relational knowledge not expressible in attribute-based systems
  - Done top-down (refine a general rule)/bottom-up (invert deductive process)
  - Can generate new predicates to make concise theories