23.01.2024

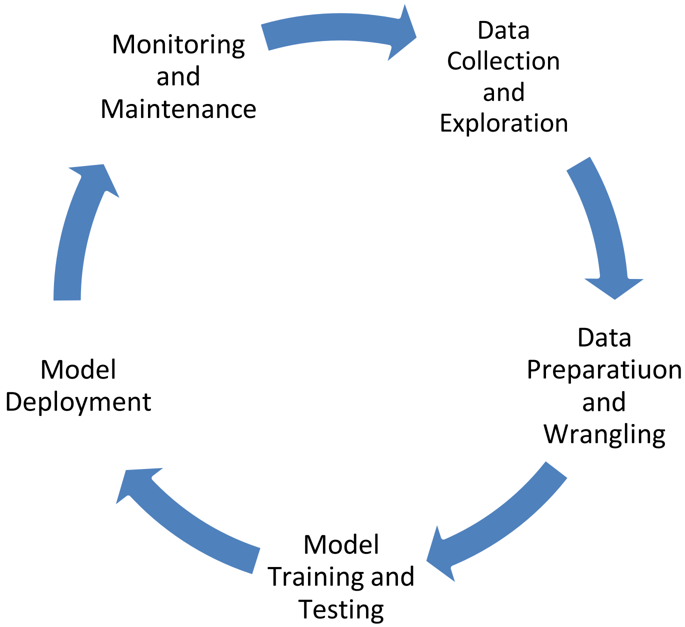
BITS F464: Machine Learning

SYMBOLIC ML: CONCEPT LEARNING

Chittaranjan Hota, Sr. Professor
Dept. of Computer Sc. and Information Systems
hota@hyderabad.bits-pilani.ac.in

Recap

- ML Frameworks
- Scikitlearn
- Pandas
- Matplotlib
- SciPy
- TensorFlow
- Keras
- PyTorch

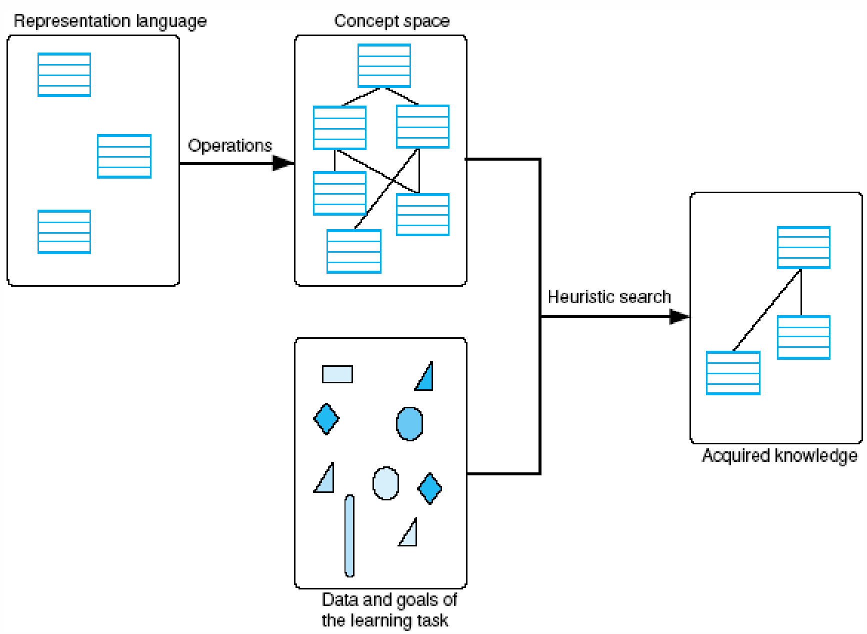


(Machine Learning Life Cycle)

Approaches to Machine Learning

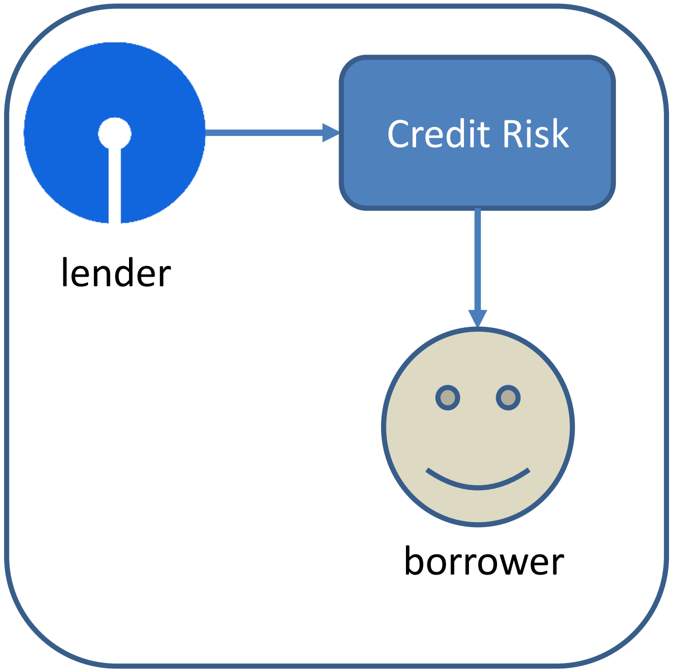
- Symbolic Learning
  - Knowledge Representation: Symbols and Rules
  - Learning Approach: Manipulation of Symbols
  - Interpretable and Transparent
  - Small Training data
- Connectionist Learning
  - Knowledge Representation: Distributed layers
  - Learning Approach: Weight adjustment
  - Highly non-linear relations
  - Large amounts of training data
- Probabilistic Learning
  - Ability to take into account uncertainty
- Evolutionary Learning

Concept Learning: A General Model



- Inferring a Boolean-valued function/hypothesis from labeled training examples. c:X → {0, 1}
- h: X → { 0, 1 }, the goal is to find h such that h(x) = c(x) for all x in X.

How Concept Learning Infers a function?



(Credit Risk Assessment)

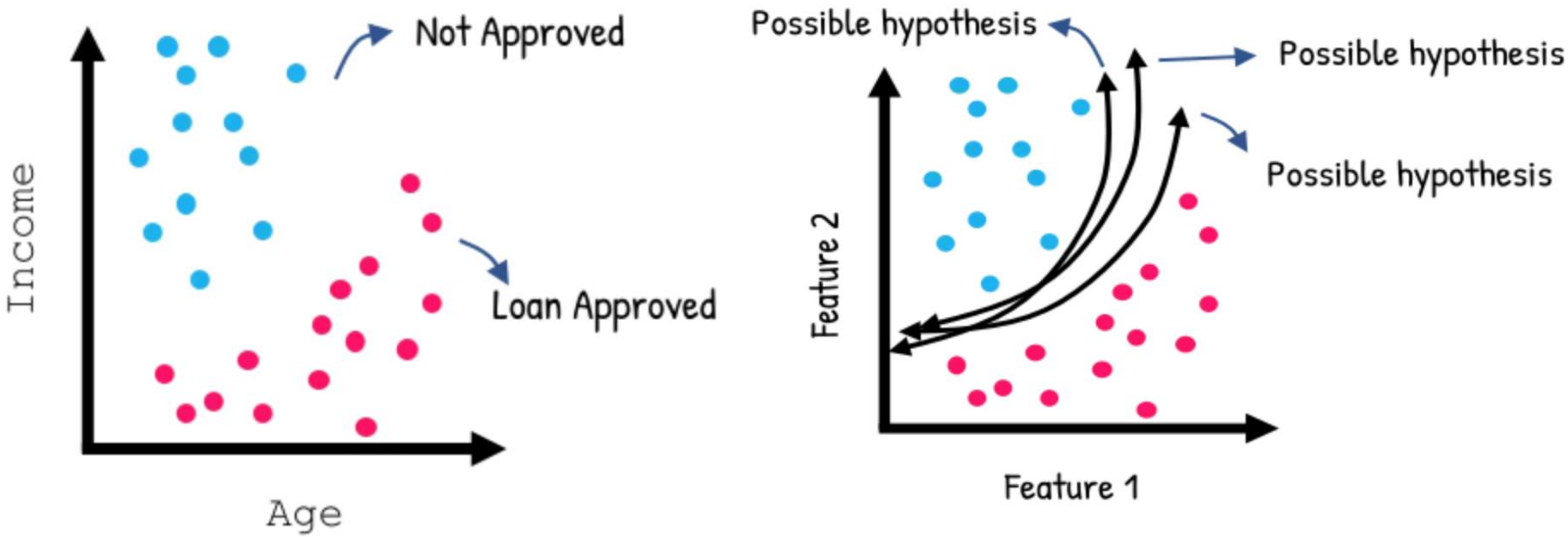
Gender	Age Group	Income	Dependents	Loan Amount	Award Loan?
Male	<18	3000	0	120	No
Female	35 - 50	2350	0	100	No
Female	35 - 50	4500	1	90	Yes
Male	18 - 25	6000	2	120	Yes

Negative Example

Positive Example

Img. Source: https://www.analyticsvidhya.com/

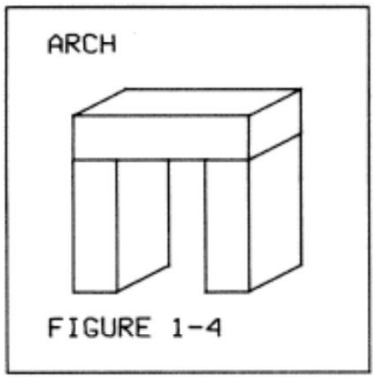
Credit Risk Assessment Continued…

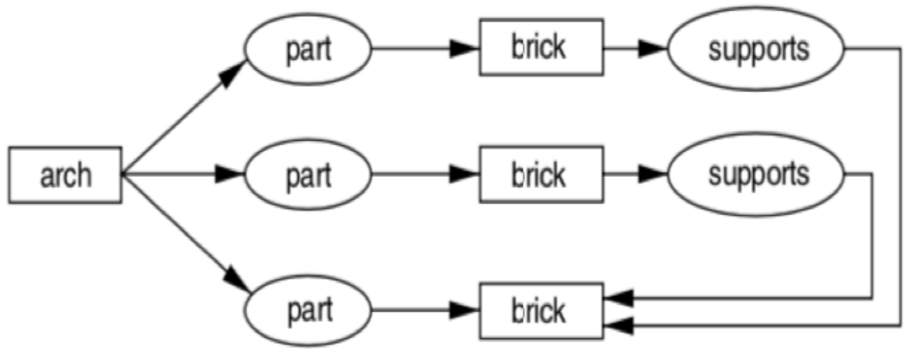


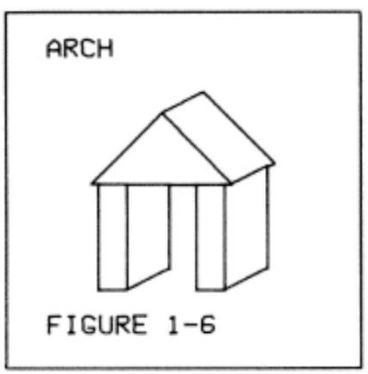
Source: https://www.analyticsvidhya.com/

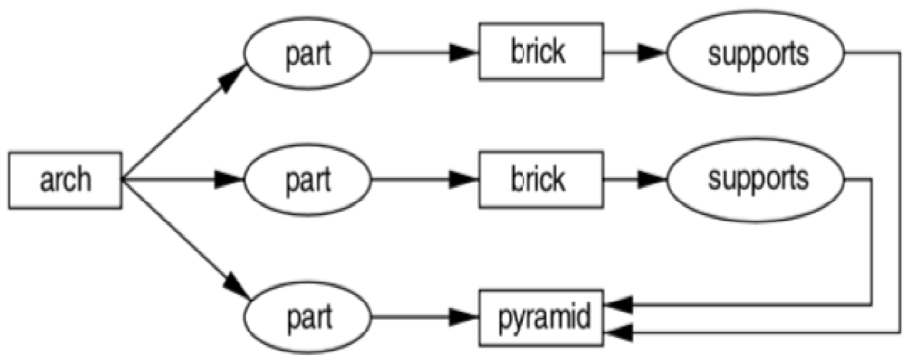
Winston's Program: Inductive Learning

(Learning Concept "Arch")





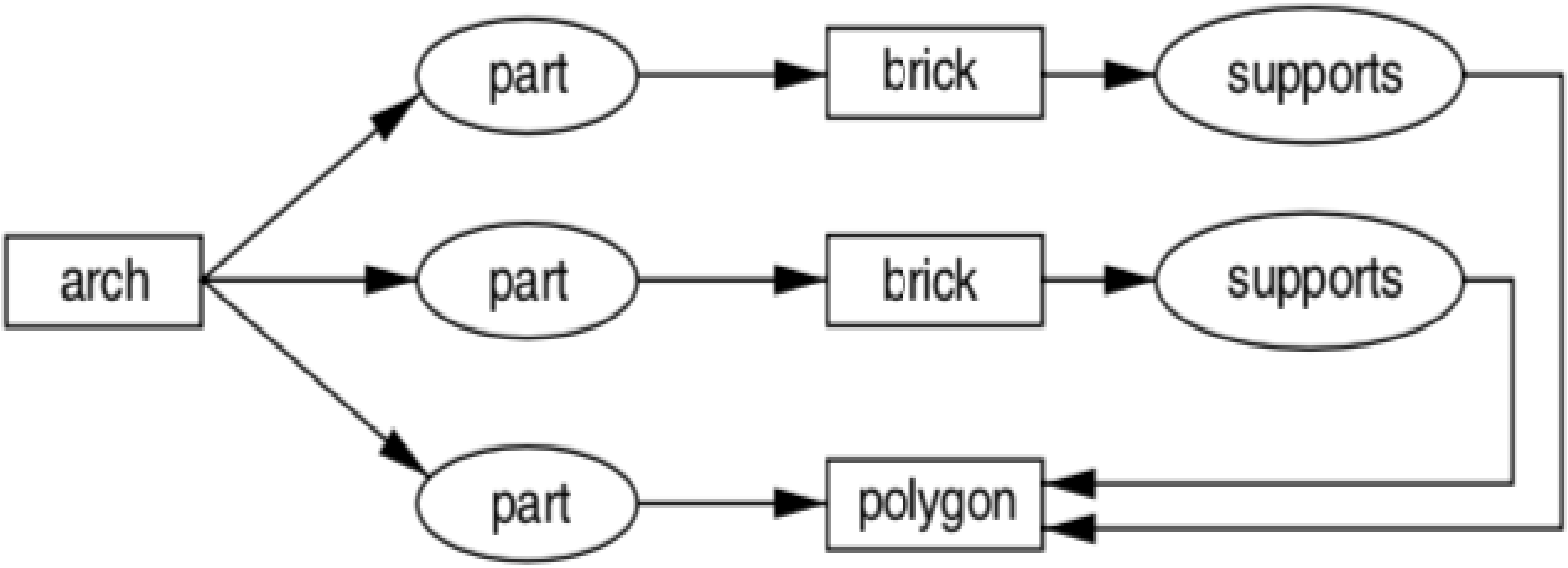




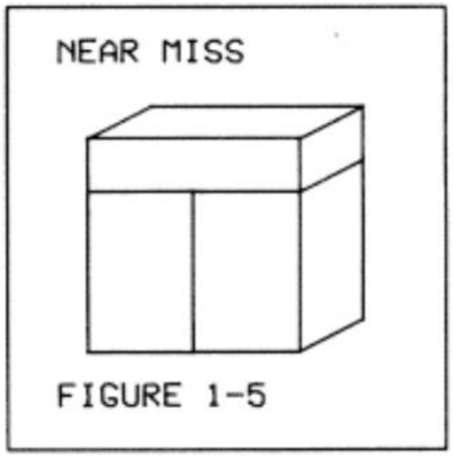
Patrick Winston: Learning Structural Descriptions from Examples, PhD Thesis (1970, MIT)

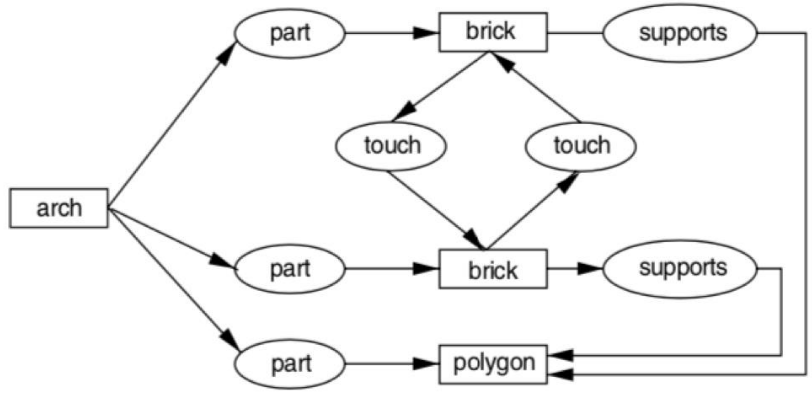
Generalization using background knowledge

- What background knowledge learner can use here to generalize?

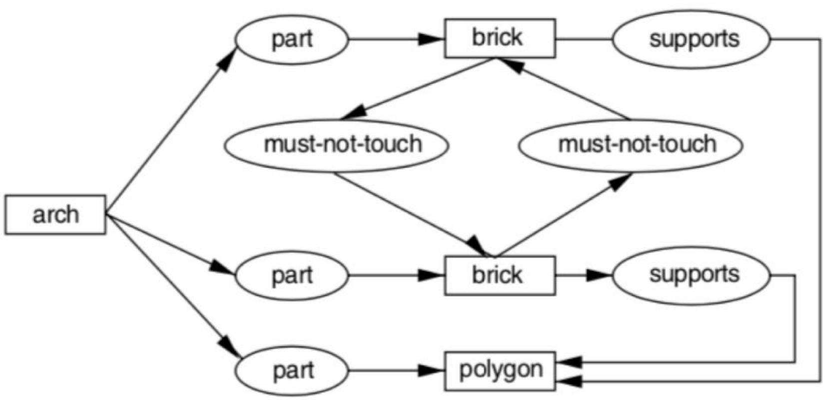


Description of a “near miss” & Specialization





How important is order of examples ?

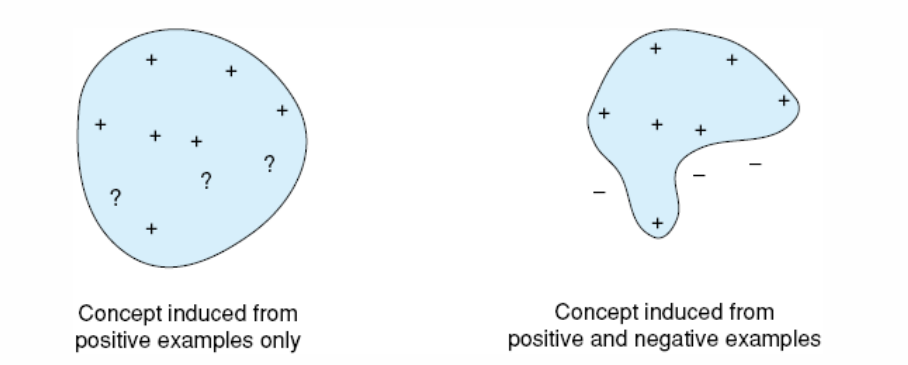


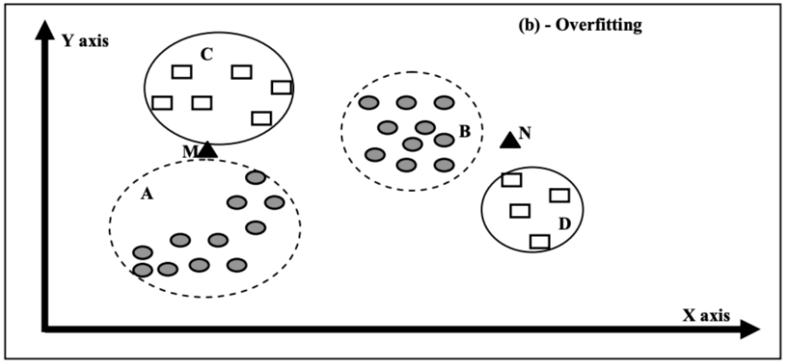
Hill Climbing

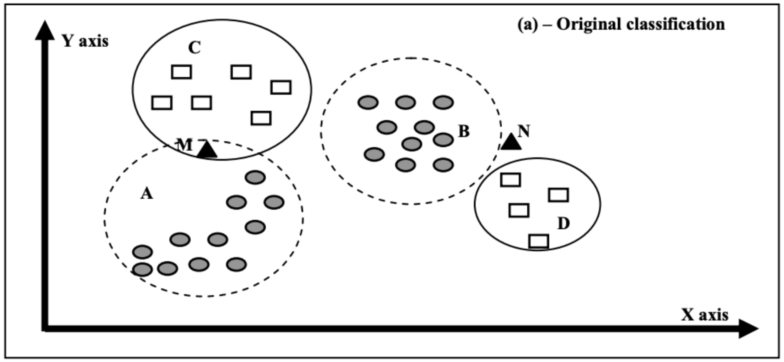
More at: https://dspace.mit.edu/bitstream/handle/1721.1/6884/AITR-231.pdf?sequence=2

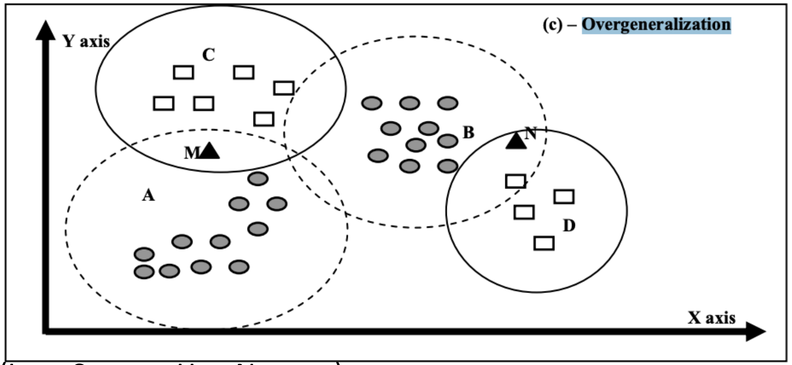
The Role of Negative Examples

- Negative instances prevent overgeneralization by forcing the learner to specialize concepts in order to exclude negative instances.





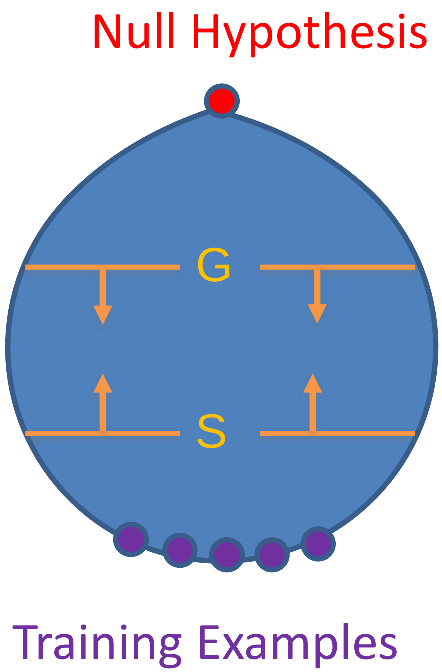


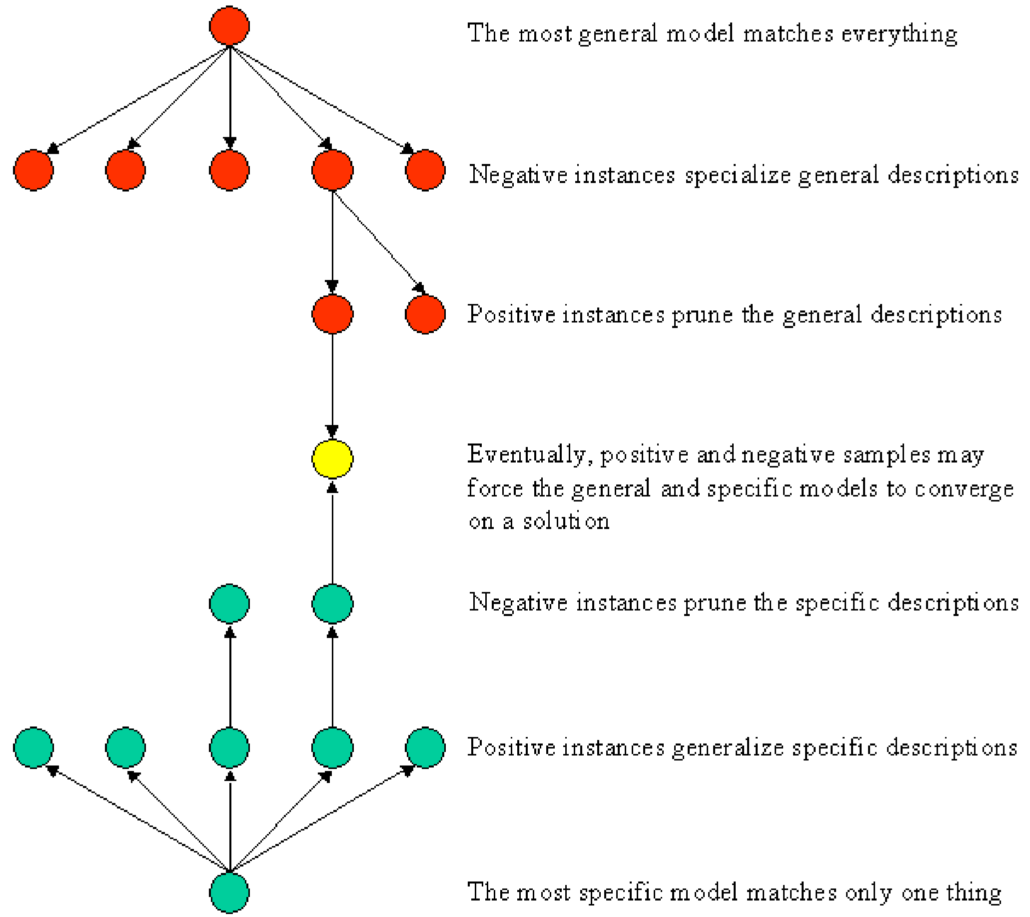


(Img. Source: Huy Nguyen)

Version Space as a Search

- Inductive learning as a Search through the Concept Space





Candidate Elimination (Mitchell, PhD, Stanford)

- Initialize G to contain one element: the most general description (all features are variables).
- Initialize S to empty.
- Accept a new training example.

(Tom Mitchell, CMU)

- Process Positive Examples:
- Remove from G any hypothesis that do not cover the example.
- Generalize S as little as possible so that the new training example is covered.
- Remove from S all elements that cover negative examples.

Algorithm continued…

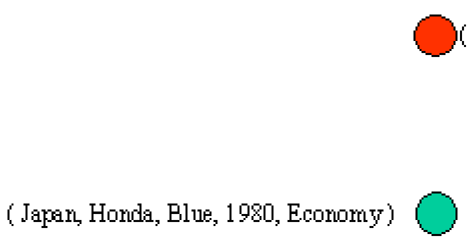
- Process Negative Examples:
- Remove from S any descriptions that cover the negative example.
- Specialize G as little as possible so that the negative example is not covered.
- Remove from G all elements that do not cover the positive examples.
- Continue processing new training examples, until one of the following occurs:
- Either S or G become empty, there are no consistent hypotheses over the training space. Stop.
- S and G are both singleton sets.
  - if they are identical, output their value and stop.
  - if they are different, the training cases were inconsistent. Output this result and stop.
- No more training examples. G has several hypotheses.

Example

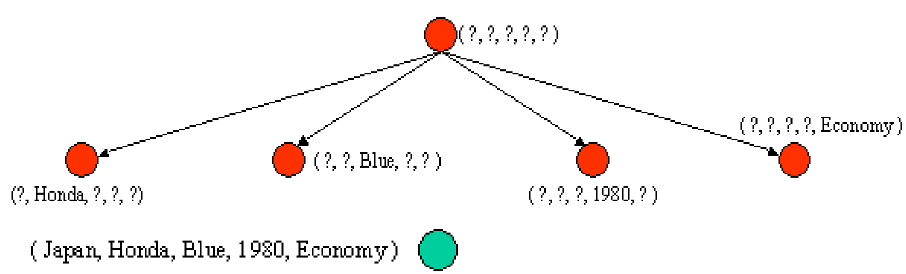
Learning the concept of "Japanese Economy Car"

Origin	Manufacturer	Color	Decade	Type	Example Type
Japan	Honda	Blue	1980	Economy	Positive
Japan	Toyota	Green	1970	Sports	Negative
Japan	Toyota	Blue	1990	Economy	Positive
USA	Chrysler	Red	1980	Economy	Negative
Japan	Honda	White	1980	Economy	Positive

1. +ve:(Japan,Honda,Blue,1980,Economy)

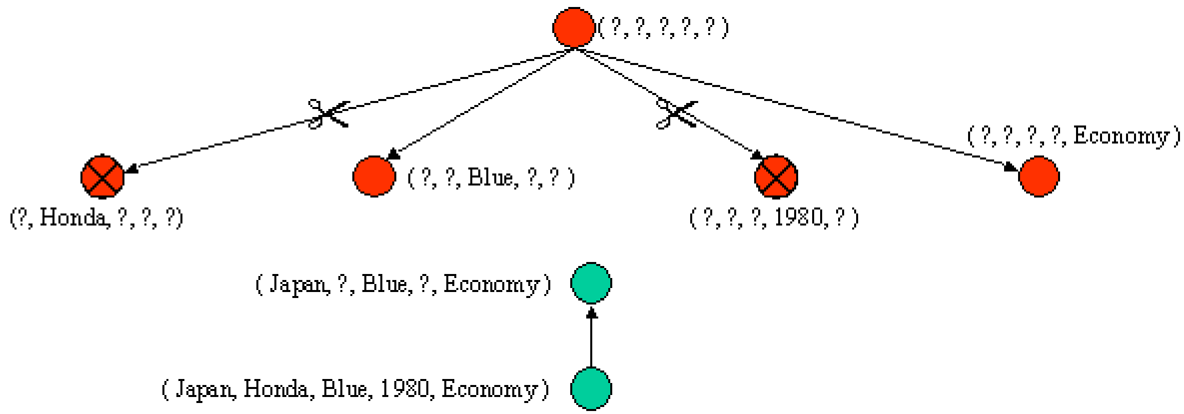


2. –ve:(Japan,Toyota,Green,1970,Sports)

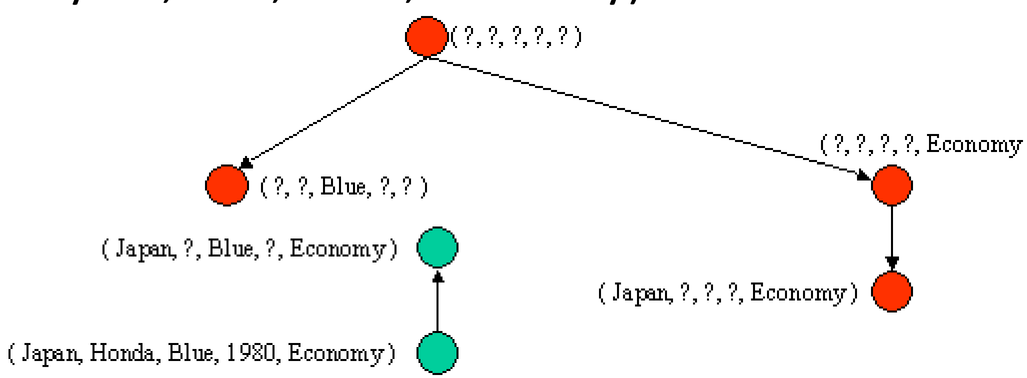


Example continued…

3. +ve : (Japan, Toyota, Blue, 1990, Economy)



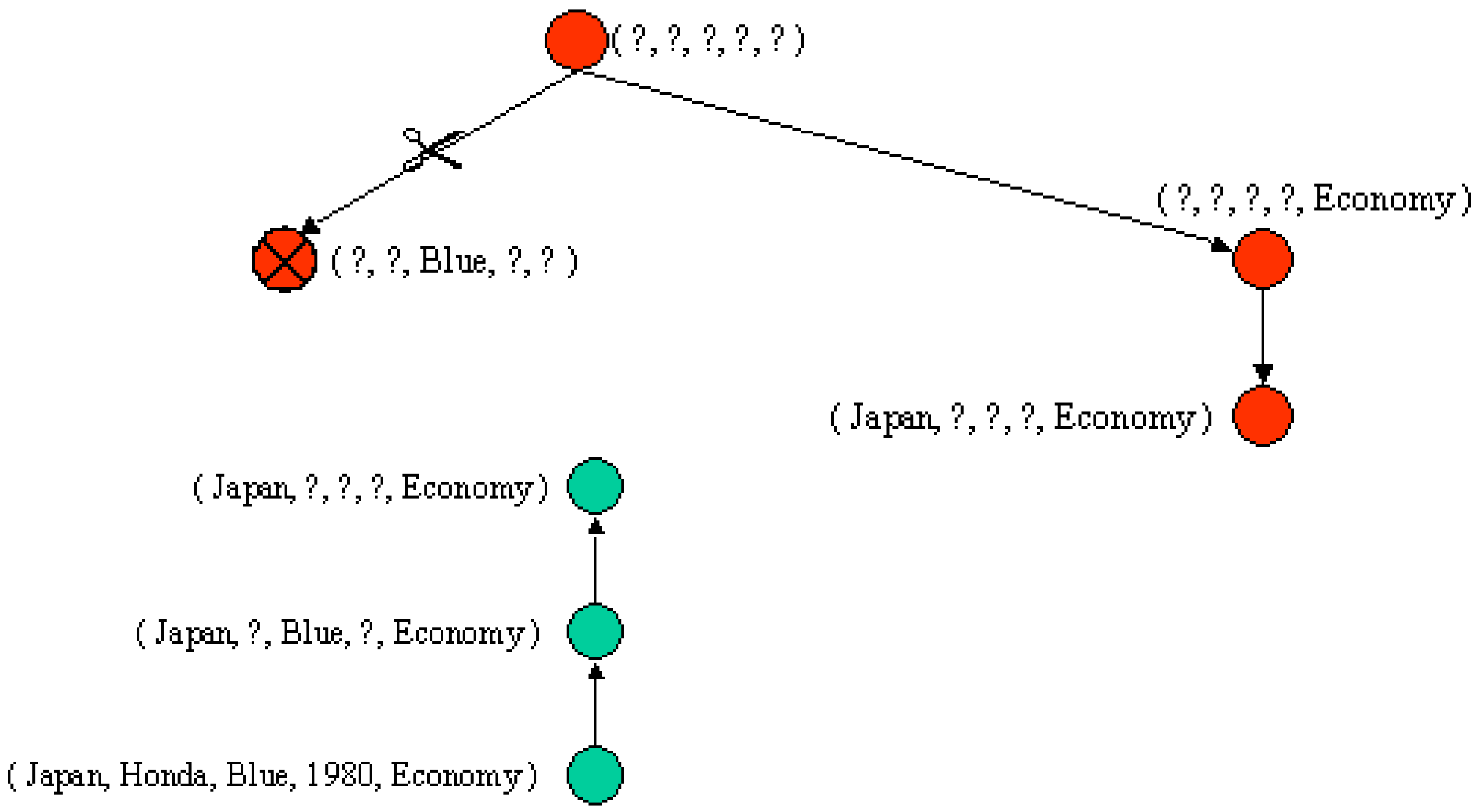
4. –ve : (USA, Chrysler, Red, 1980, Economy)



Img. Source: https://www2.cs.uregina.ca/~dbd/cs831/notes/ml/vspace/vs_prob1.html

Example continued…

5. +ve : (Japan, Honda, White, 1980, Economy)



How important are training examples?

Origin	*Manufacturer*	*Color*	*Decade*	*Type*	*Example Type*
Japan	Honda	Blue	1980	Economy	Positive
Japan	Toyota	Green	1970	Sports	Negative
Japan	Toyota	Blue	1990	Economy	Positive
USA	Chrysler	Red	1980	Economy	Negative
Japan	Honda	White	1980	Economy	Positive
Japan	Toyota	Green	1980	Economy	Positive
Japan	Honda	Red	1990	Economy	Negative

Consistent with version space, and hence G: (Japan,?,?,?, Economy), and S: (Japan, ?, ?, ?, Economy) i.e. No change.

In-consistent with version space and hence, G: ϕ, and S: ϕ (no concept)

Biased Hypothesis Space

- Candidate elimination will converge towards the target concept, provided:
  - Accurate training examples are available to the learner
  - Initial hypothesis space contains the target concept
- If target concept is not present, then it is Biased.

Example	*Sky*	*AirTemp*	*Humidity*	*Wind*	*Water*	*Forecast*	*EnjoySport*
1	Sunny	Warm	Normal	Strong	Cool	Change	Yes
2	Cloudy	Warm	Normal	Strong	Cool	Change	Yes
3	Rainy	Warm	Normal	Strong	Cool	Change	No

ϕ

S2 : (?, Warm, Normal, Strong, Cool, Change)

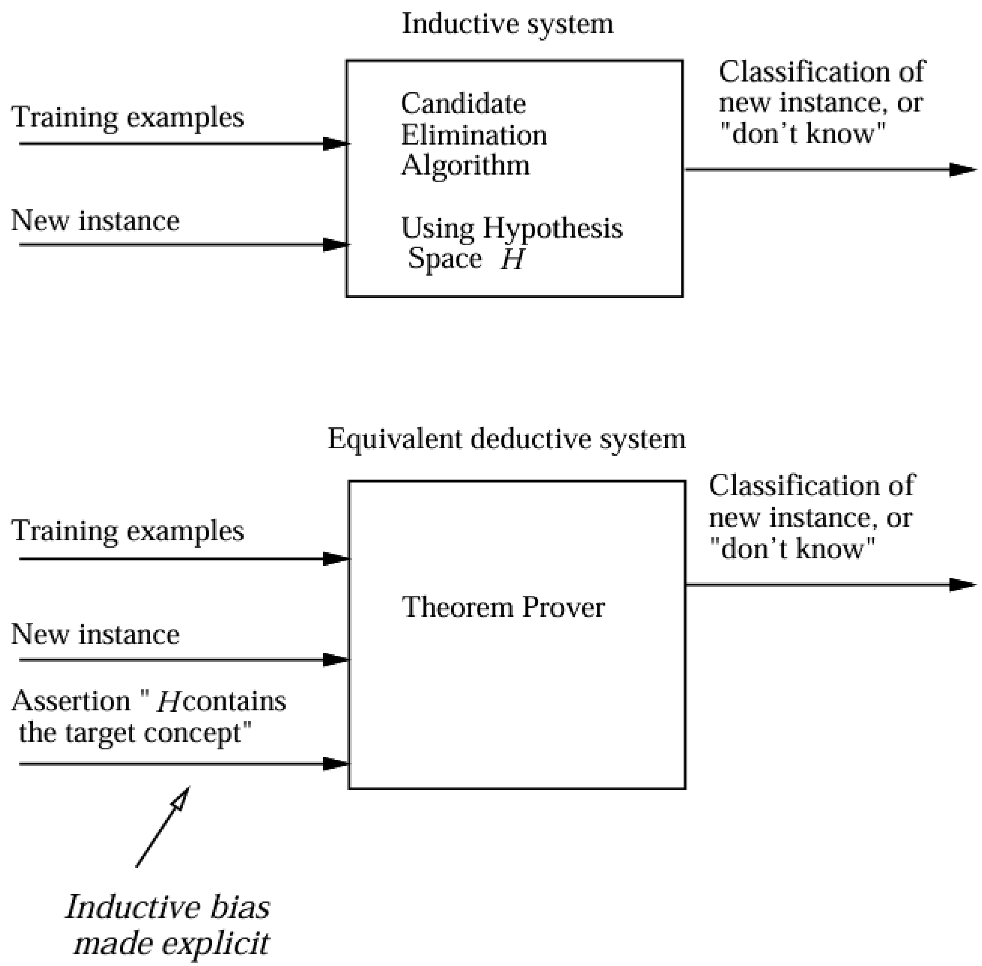
S1 : (Sunny, Warm, Normal, Strong, Cool, Change)

The problem is that we have biased the learner to consider only <u>conjunctive</u> hypotheses.

Inductive Bias in Concept Learning

- Idea: Choose H that expresses every teachable concept (i.e. H is a powerset of X).

- Consider H’ as disjunction, conjunction, and negation over previous H.



Quiz for you…

- Which of the following is NOT a knowledge representation scheme for Symbolic ML?
  - Propositional and Predicate Logic
  - Semantic Networks
  - Bayesian Networks ✓
- Candidate elimination algorithm takes on training examples and searches what to find out a target concept?
  - Concept space
  - Version space ✓
  - Decision Tree
- Which one does not contribute to Overgeneralization in ML?
  - Sufficient training data ✓
  - Imbalanced training data
  - Biased training data
  - Overfitting

Thank you!