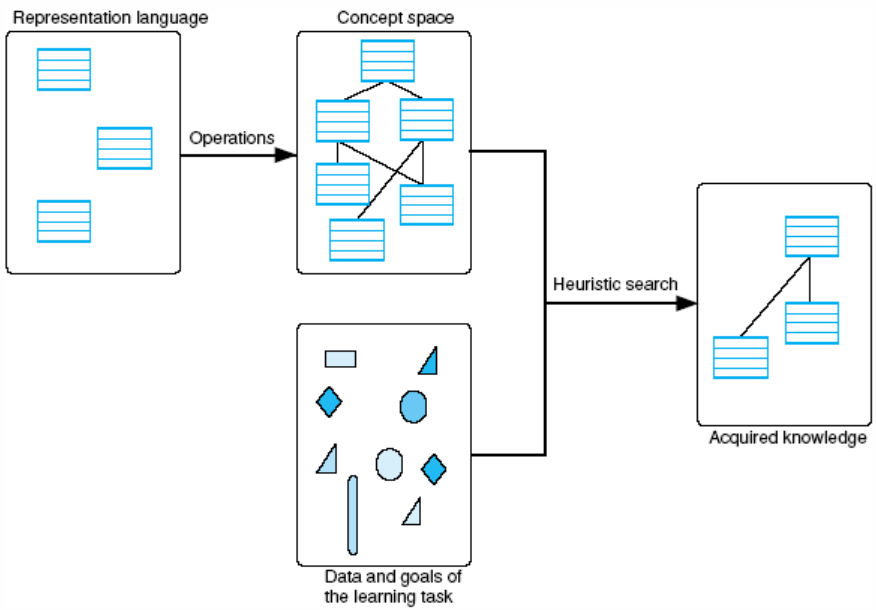
21.08.2024

BITS F464: Machine Learning (1st Sem 2024-25)

SUPERVISED LEARNING-I (CONCEPT LEARNING)

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

Concept Learning: A General Model



(BITS, Hyd Tennis court)

Concept: Stop

Rule: If the traffic light is red, then stop.

Concept: Bird

Rule: If an animal has feathers, can fly, and lays eggs, then it is a bird.

Concept: Play tennis

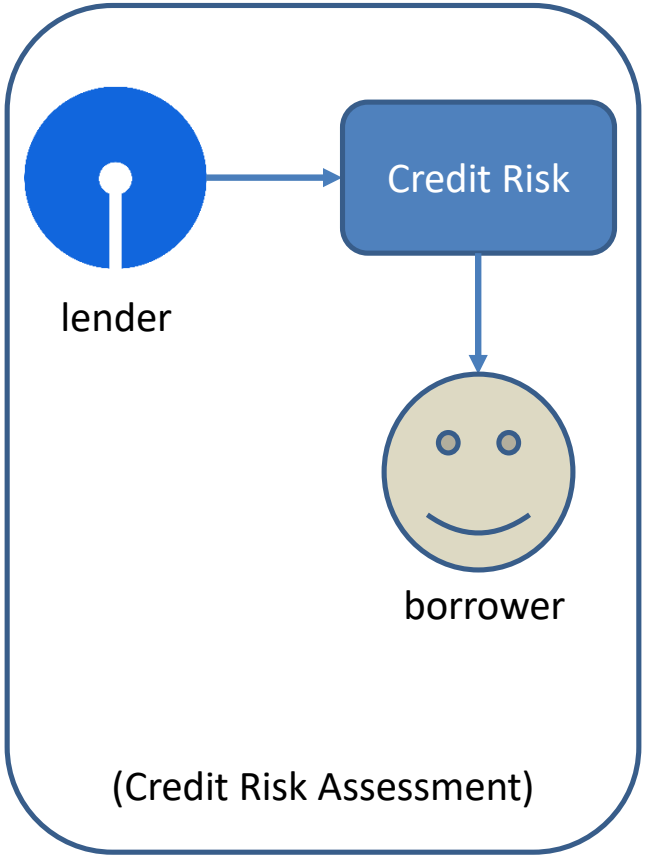
Rule: If the weather is sunny, the humidity is normal, and the wind is not strong, then play tennis.

Concept: Flu

Rule: If a person has a fever, cough, and body aches, then they have the flu.

Question: What will you learn given 100 spam emails and 200 legitimate emails?

How Concept Learning Infers a function?

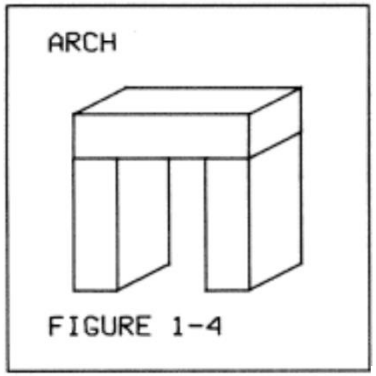


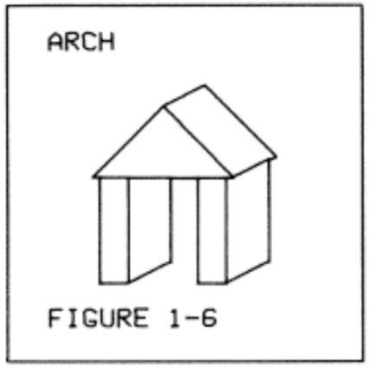
Applicant Name	Income (Annual)	Credit Score	Debt-to-Income Ratio (%)	Employment Duration (Years)	Payment History	Savings Balance	Loan Inquiries (Last 6 months)	Bankruptcy (Last 5 years)	Credit Risk
Raj	50,000	720	30%	5	No late payments	15,000	1	No	Low Risk
Sham	30,000	650	40%	3	1 late payment	5,000	2	No	Medium Risk
Nitin	75,000	780	20%	10	No late payments	25,000	0	No	Low Risk
Bikash	40,000	600	50%	2	Multiple late payments	2,000	3	No	High Risk

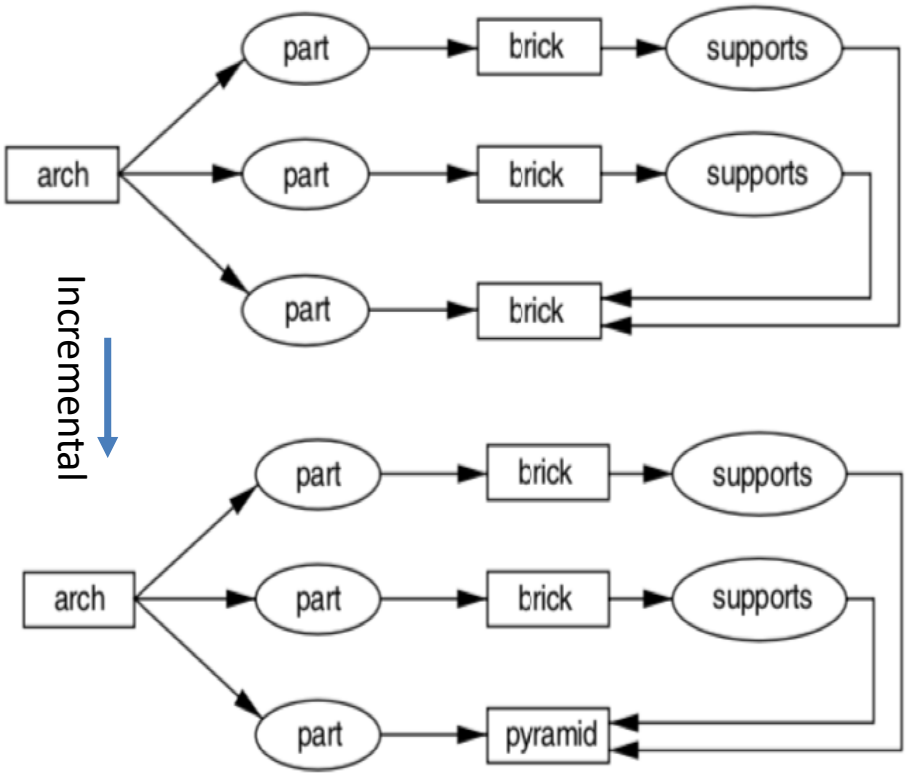
Question: What function will you learn here?

Winston's Program: Inductive Learning

(Learning Concept "Arch")







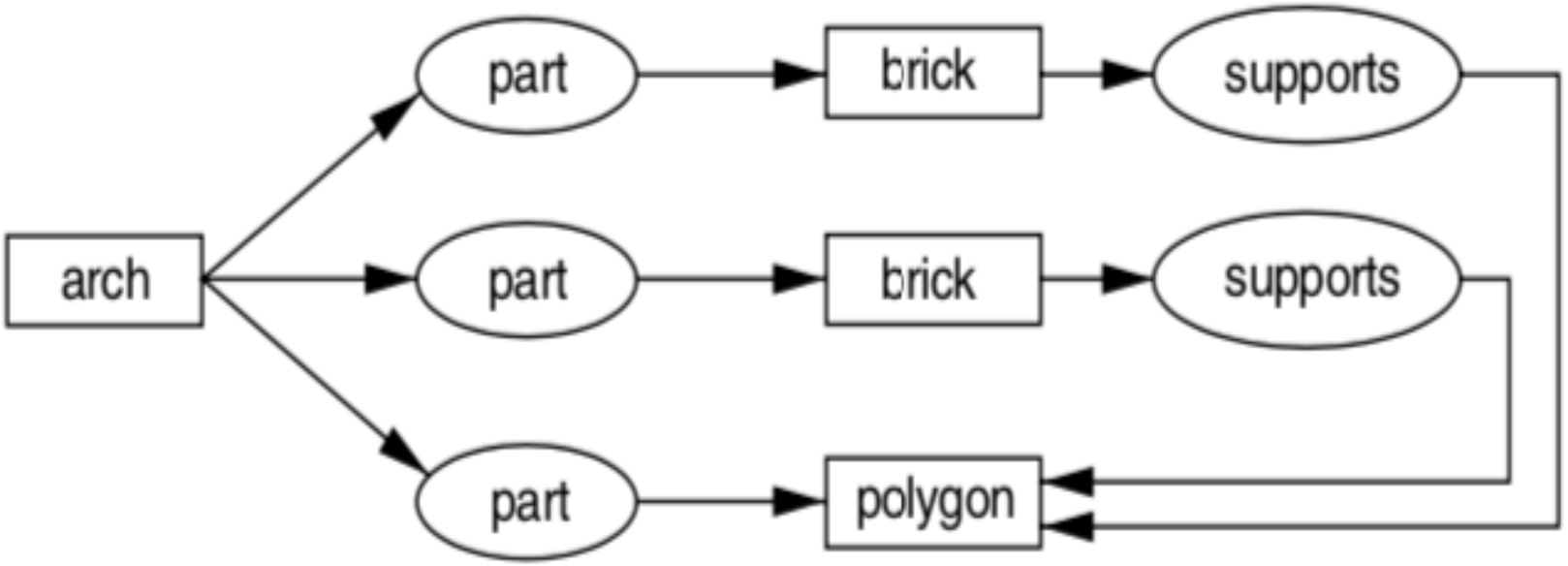
Director of MIT's AI Lab: 1972-97

Generalization from Specifics

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

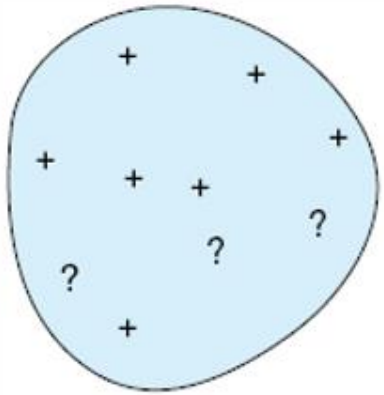
Generalization using background knowledge

- What background knowledge learner can use here to generalize?

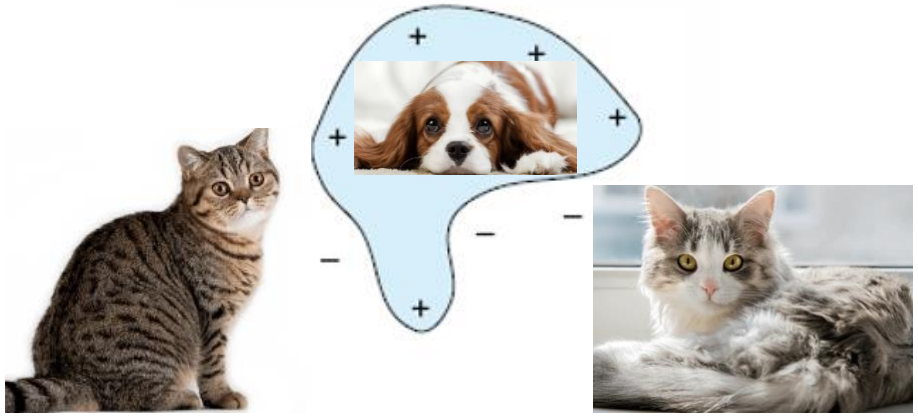


The Role of Negative Examples

- Negative instances prevent overgeneralization by forcing the learner to specialize concepts in order to exclude negative instances.
- All small, furry, four-legged animals are dogs (by seeing a few)
- Any email containing 'Congratulations' is spam.



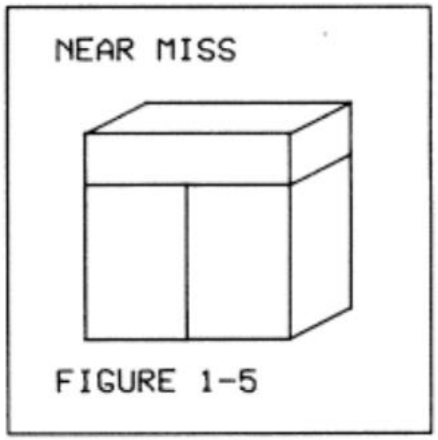
Concept induced from positive examples only

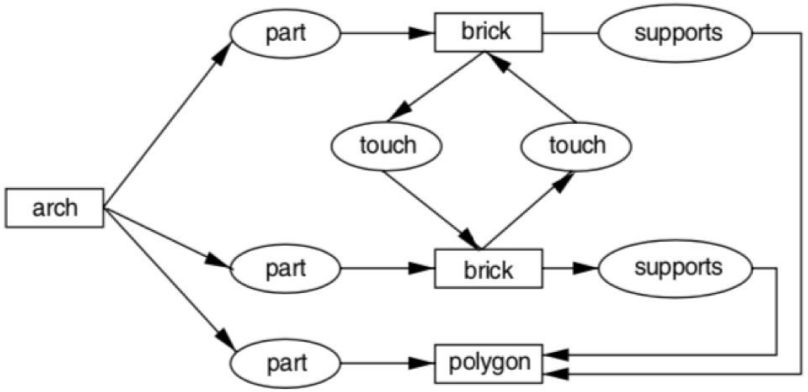


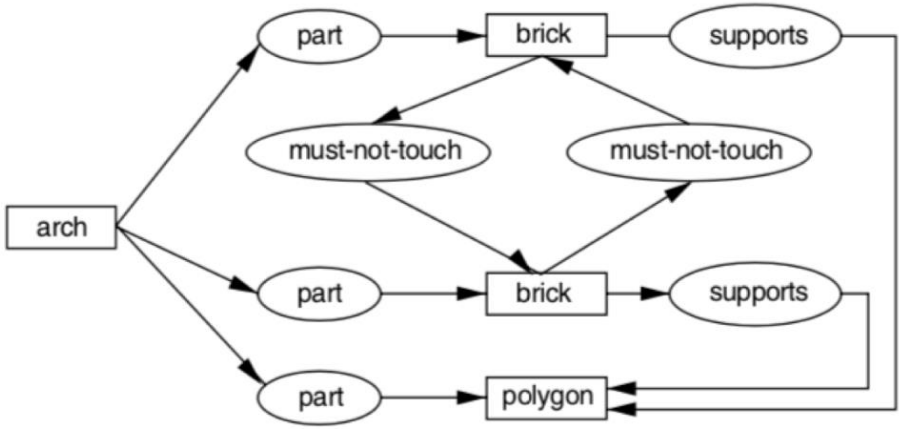
Concept induced from positive and negative examples

Question: What could be the function or rule now?

Description of a “near miss” & Specialization







Hill Climbing

How important is the order of examples?

A bird with wings, feathers, and the ability to fly. (+ve)

A penguin (a bird that cannot fly). Near miss.

A bat (which flies but is not a bird). Counter ex.

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

Recap: Winston's Program

Each example introduces a new feature while maintaining the core structure, allowing the program to generalize effectively.

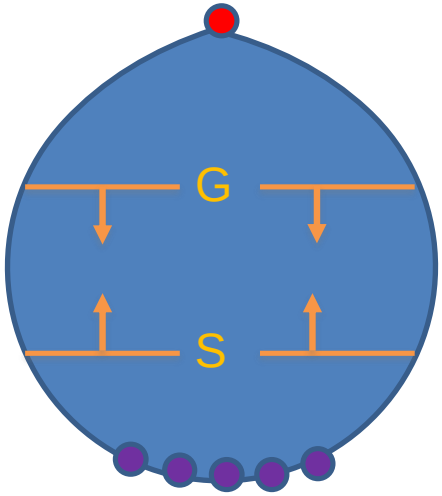
Recap: Overgeneralization

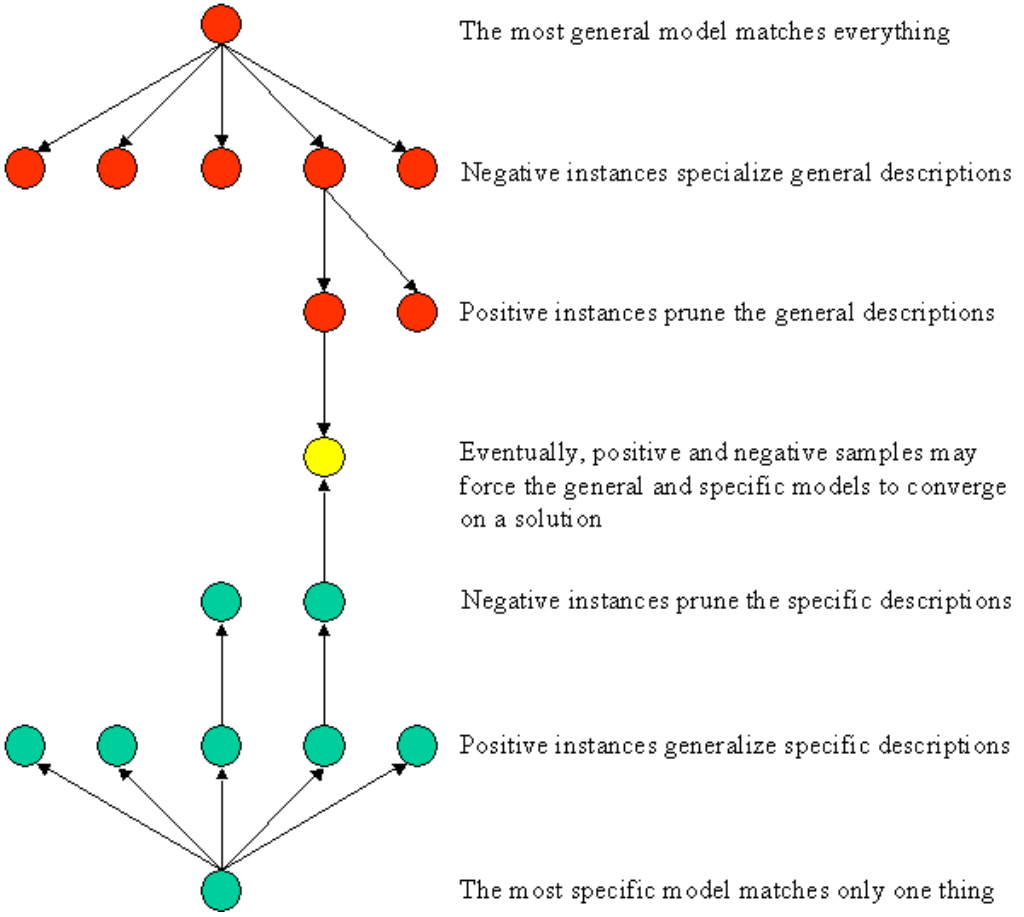
In this order, the system might mistakenly focus on the material (wood) or the presence of legs, leading to overgeneralization. → wooden table or ladder may be called as chair.

Order: A Wooden Chair, A Plastic Chair, A Wooden Table, A Bean Bag, A Stool.

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.
- 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.

(Tom Mitchell, CMU)

Algorithm continued…

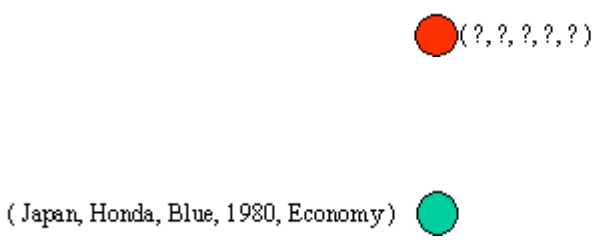
- Process Negative Examples:
- Remove from S any descriptions that cover the negative example.
- Specialize G as little as possible so that the negative example is NOT covered.
- Remove from G all elements that do NOT cover the positive examples.
- Continue processing new training examples, until one 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.

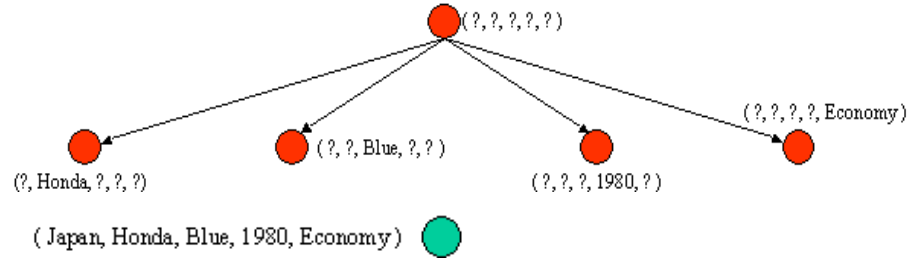
Example

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

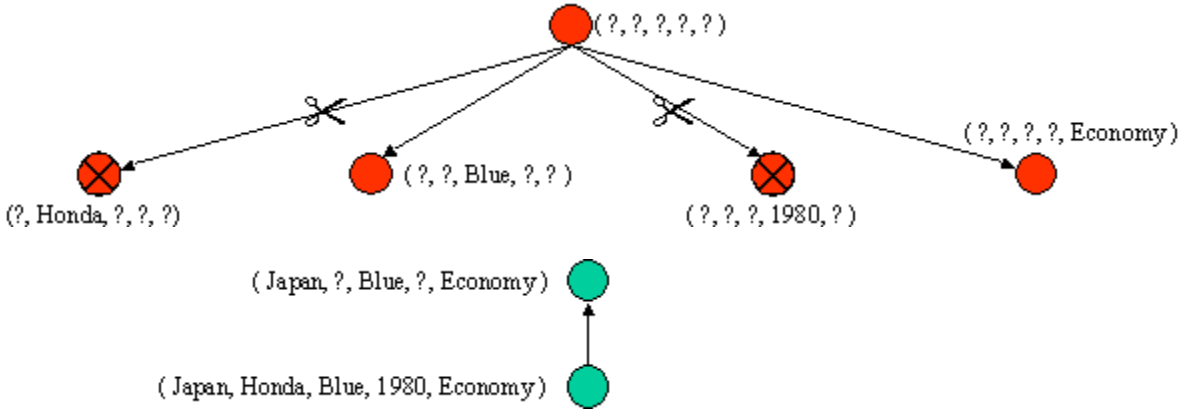
1. +ve:(Japan,Honda,Blue,1980,Economy) 2. –ve:(Japan,Toyota,Green,1970,Sports)



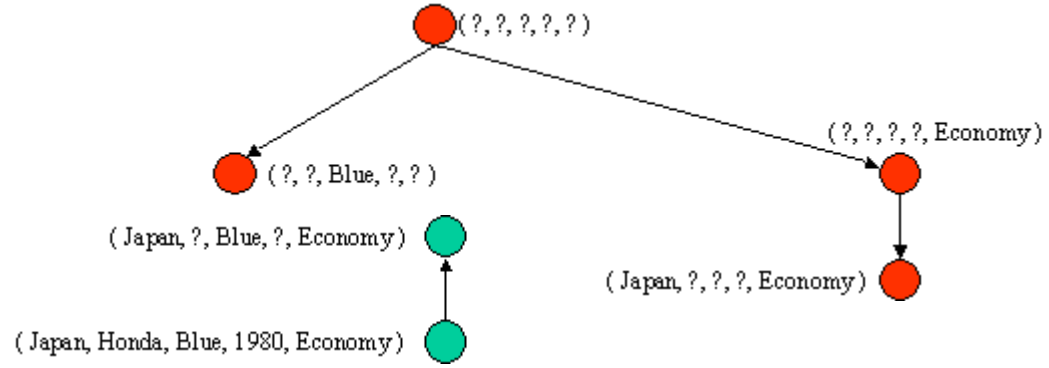


Example continued…

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



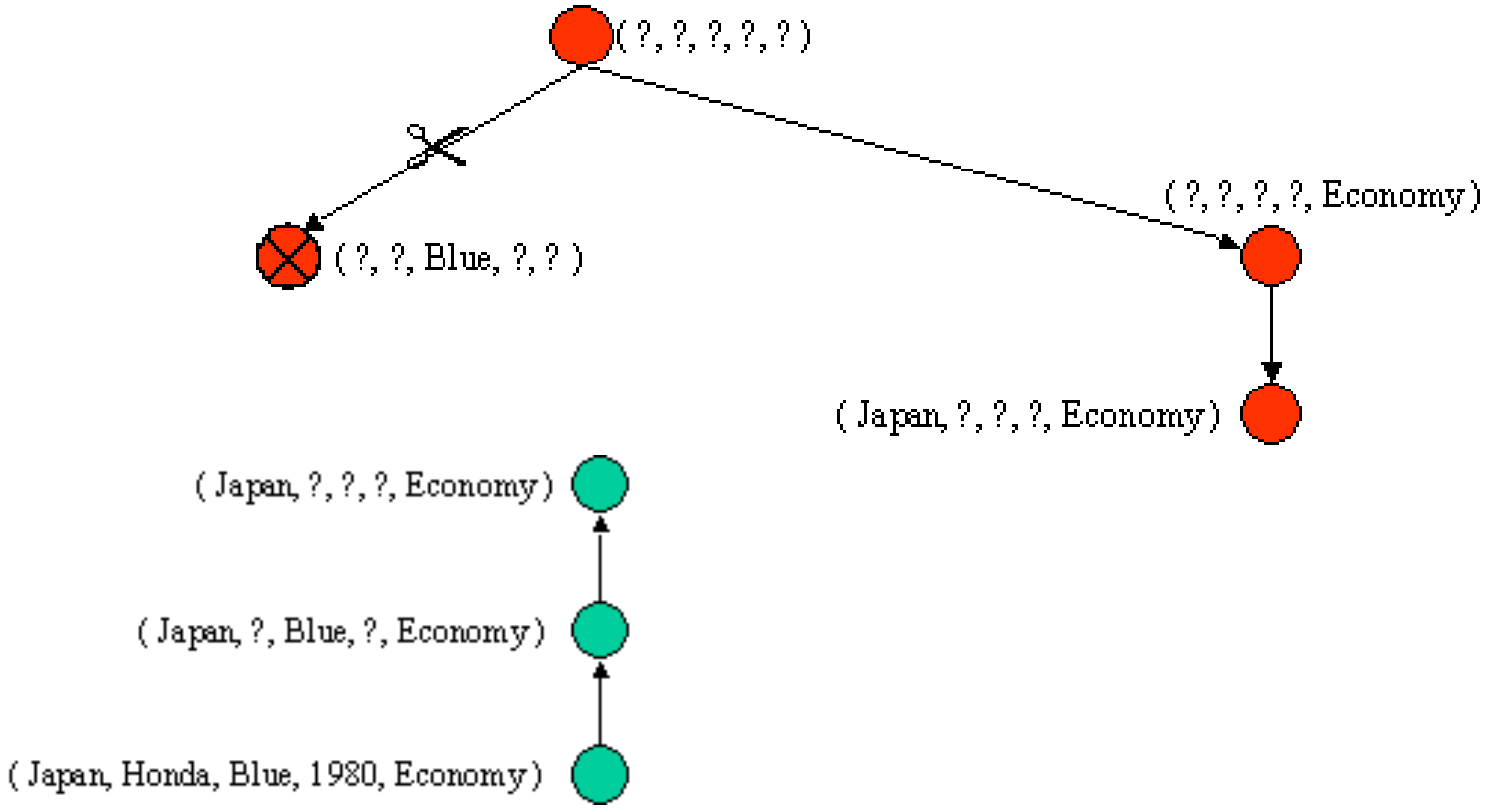
4. –ve : (USA, Chrysler, Red, 1980, Economy)



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

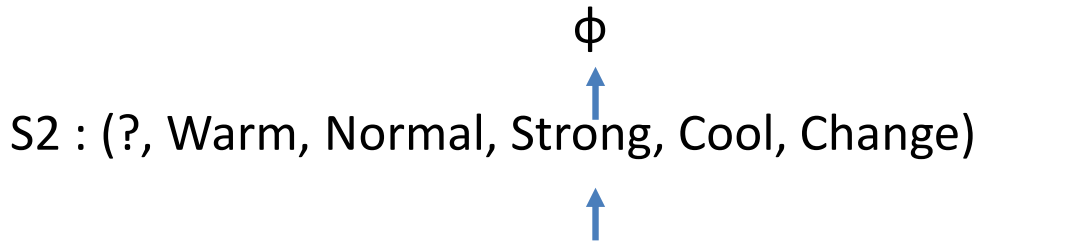
Consistent with version space, and hence G: (Japan,?,?,?, Economy), and S: (Japan, ?, ?, ?, Economy) i.e. No change.

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



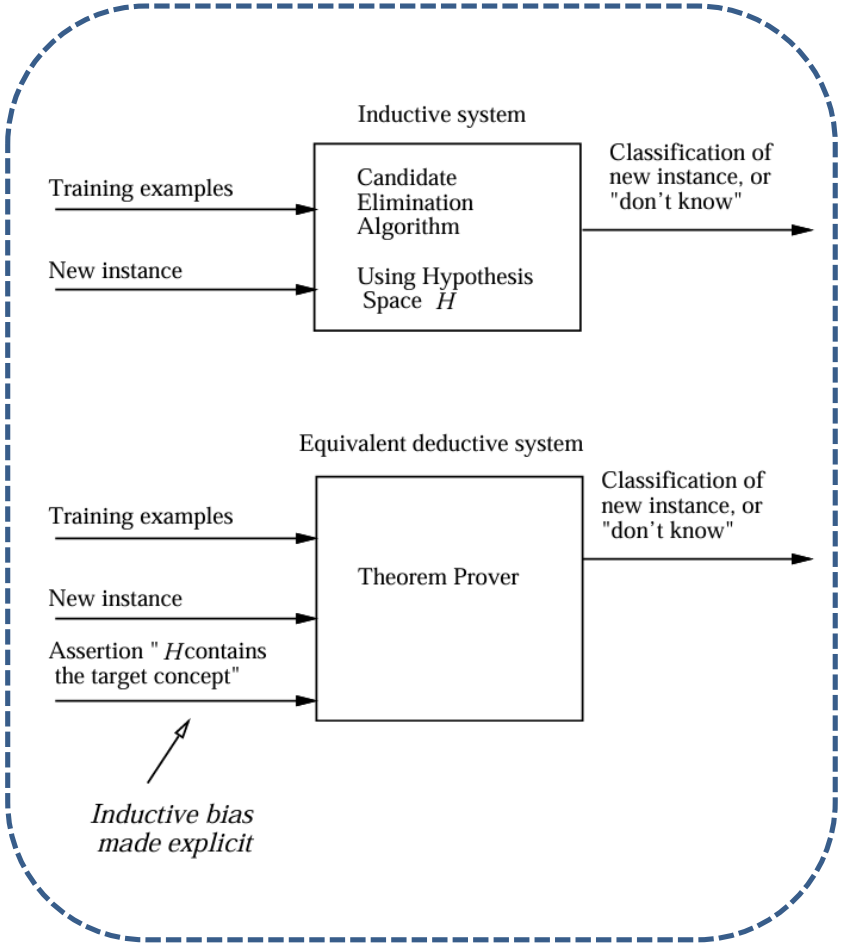
The problem is that we have biased the learner to consider only conjunctive hypotheses.

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.

Types of Inductive bias:

- Prior knowledge: dogs and cats are mammals.
- Assumptions about data distributions: classes are linearly separable of features are independent?
- Model structure: DT or ANNs



Question: What are the inductive biases when you try to apply for job?

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

Inductive bias does not support: guiding generalization, avoiding overgeneralization, ✓Not constraining hypotheses, Improving learning efficiency.

Thank You!