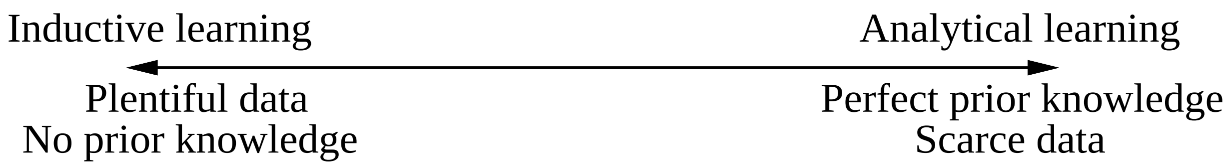
Inductive and Analytical Learning

Inductive learning	**Analytical learning**
Hypothesis fits data	Hypothesis fits domain theory
Statistical inference	Deductive inference
Requires little prior knowledge	Learns from scarce data
Syntactic inductive bias	Bias is domain theory

What We Would Like



General purpose learning method:

- No domain theory → learn as well as inductive methods
- Perfect domain theory → learn as well as Prolog-EBG
- Accomodate arbitrary and unknown errors in domain theory
- Accomodate arbitrary and unknown errors in training data

Domain theory:

$$\text{Cup} \leftarrow \text{Stable, Liftable, OpenVessel}$$
$$\text{Stable} \leftarrow \text{BottomIsFlat}$$
$$\text{Liftable} \leftarrow \text{Graspable, Light}$$
$$\text{Graspable} \leftarrow \text{HasHandle}$$
$$\text{OpenVessel} \leftarrow \text{HasConcavity, ConcavityPointsUp}$$

Training examples:

	Cups				Non-Cups					
BottomIsFlat	√	√	√	√	√	√	√			√
ConcavityPoints Up	√	√	√	√	√		√	√		
Expensive	√		√				√		√	
Fragile	√	√			√	√		√		√
HandleOnTop					√		√			
HandleOnSide	√			√					√	
HasConcavity	√	√	√	√	√		√	√	√	√
HasHandle	√			√	√		√		√	
Light	√	√	√	√	√	√	√		√	
MadeOfCeramic	√				√		√	√		
MadeOfPaper				√					√	
MadeOfStyrofoam		√	√			√				√

KBANN

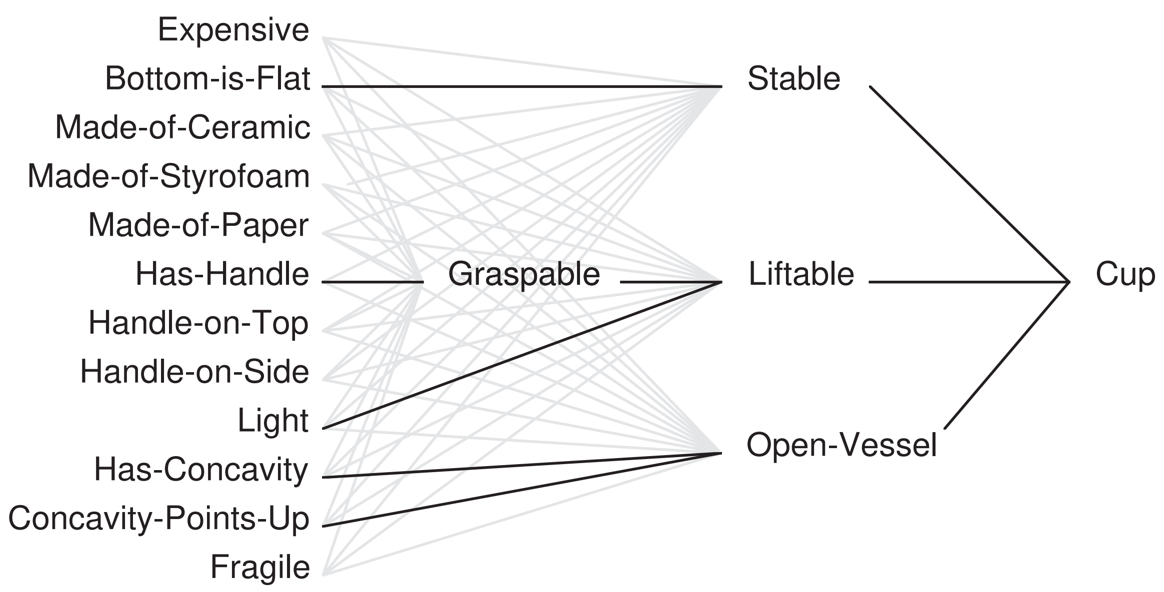
KBANN (data D, domain theory B)

1. Create a feedforward network h equivalent to B
2. Use Backprop to tune h to fit D

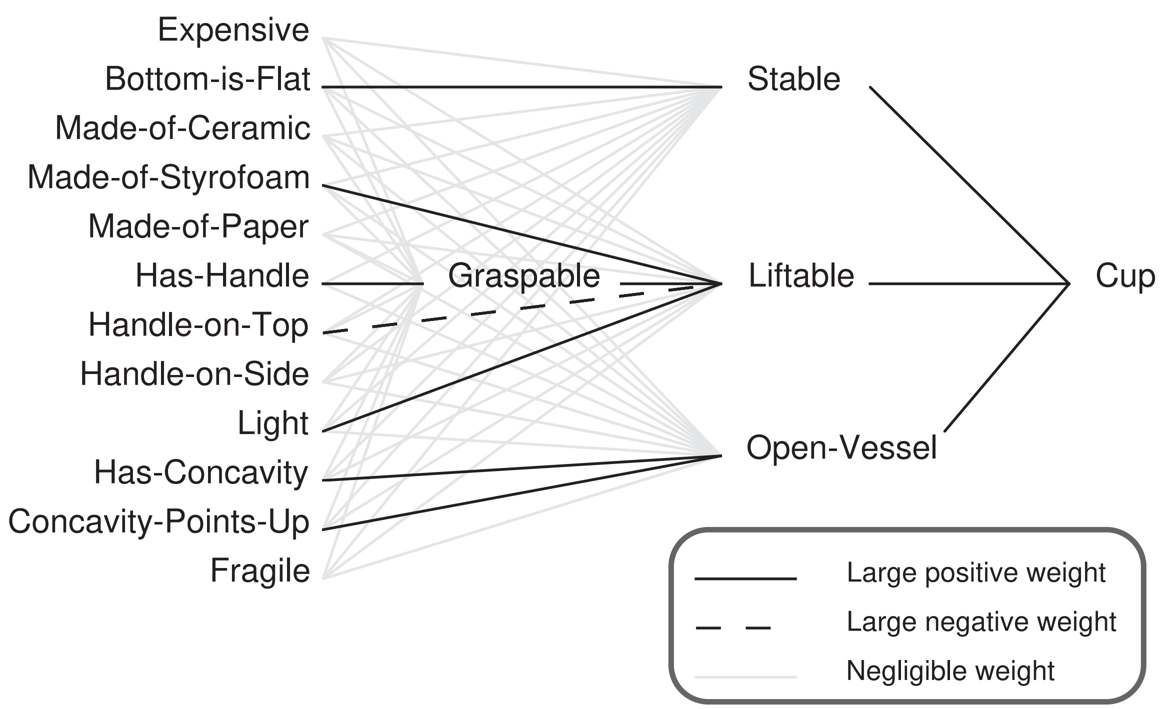
Creating a Net from the Domain Theory

(propositional theories only)

Neural Net Equivalent to Domain Theory



Result of refining the network

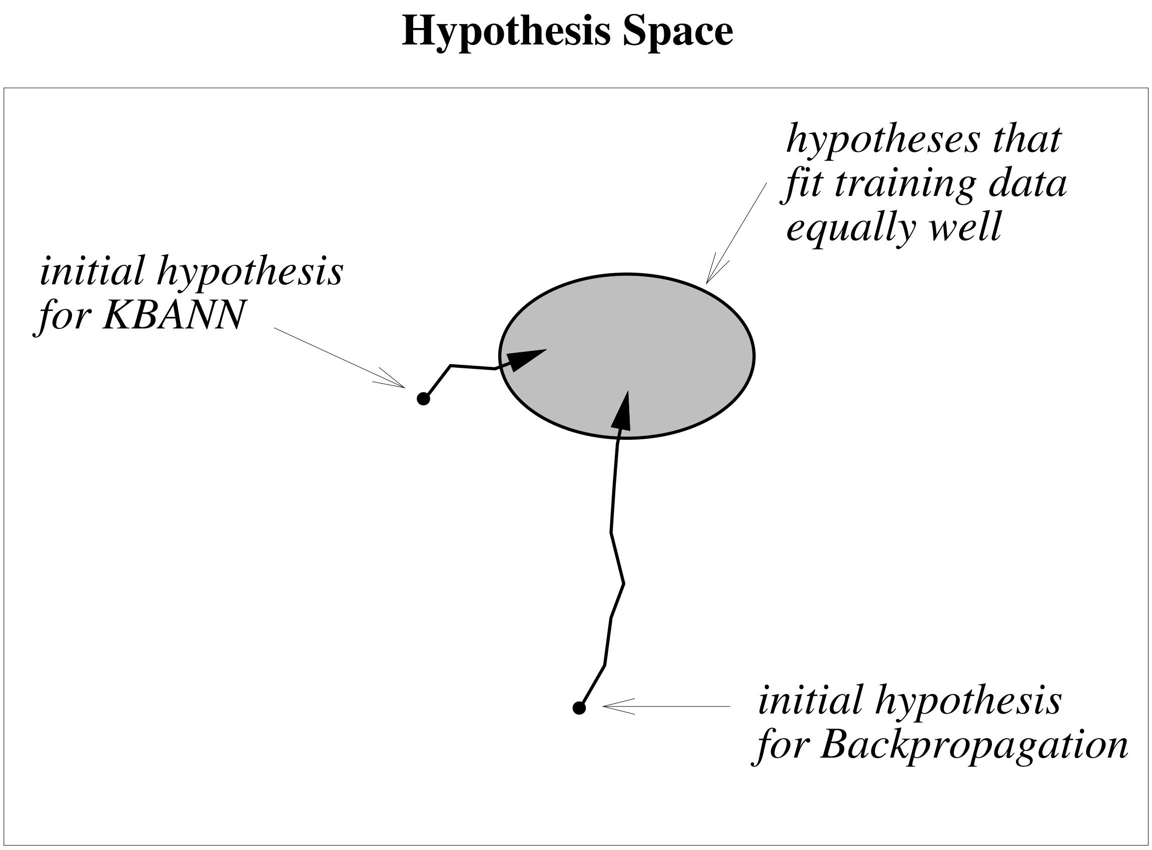


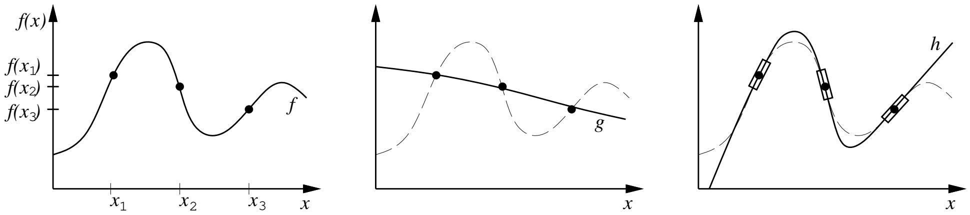
KBANN Results

Classifying promoter regions in DNA
leave one out testing:

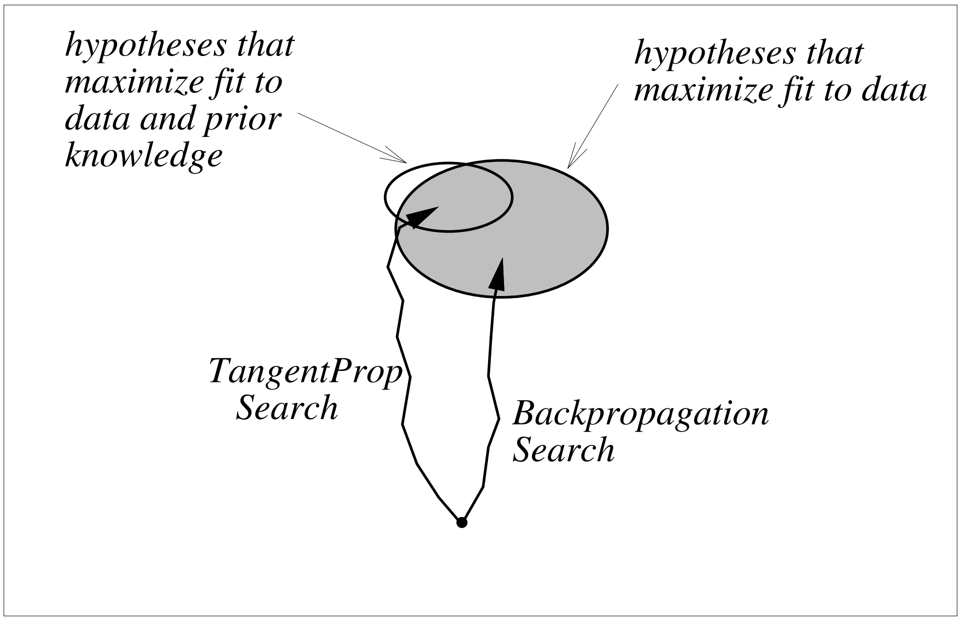
- Backpropagation: error rate 8/106
- KBANN: 4/106

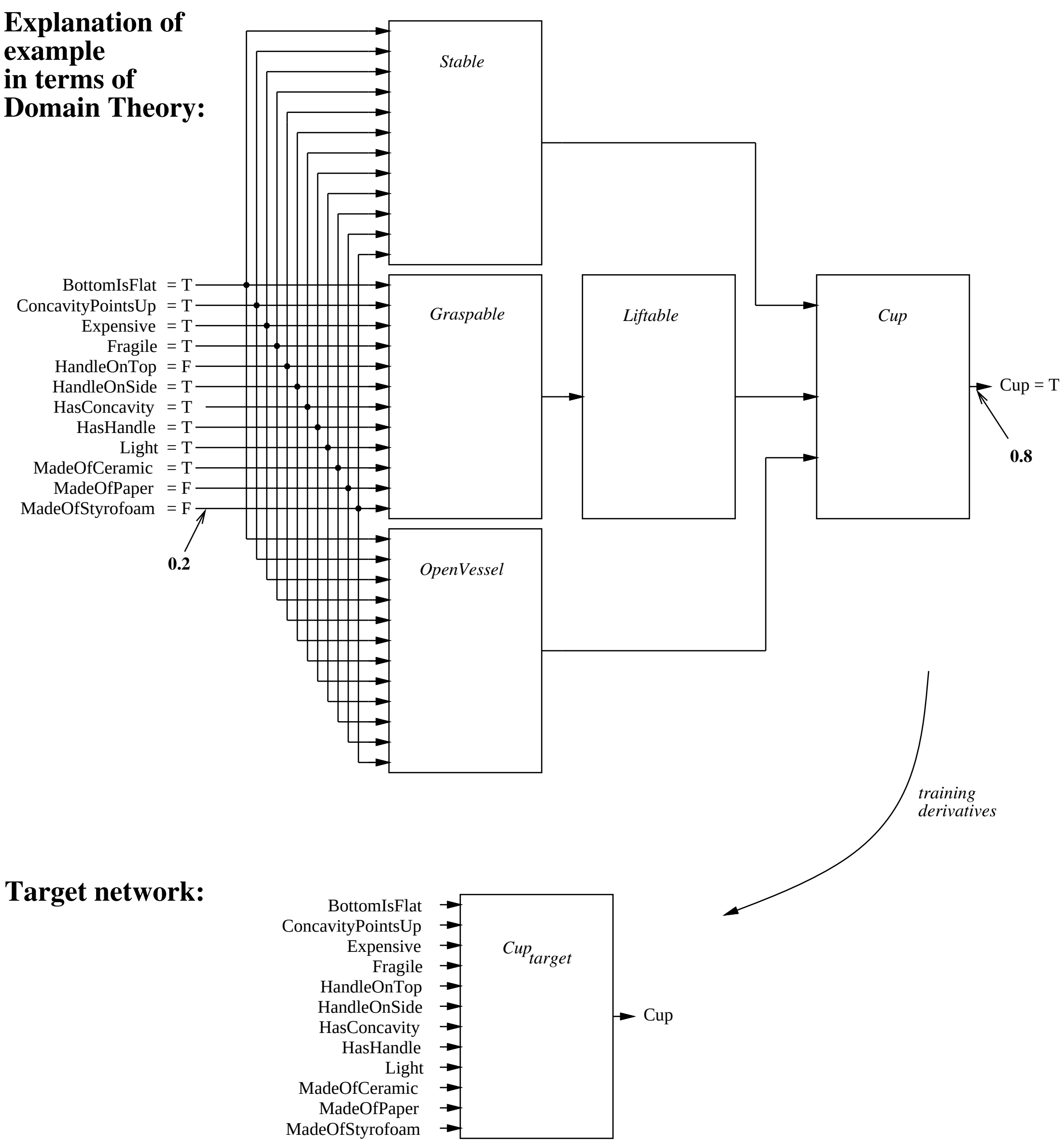
Hypothesis space search in KBANN




f(x)
f(x1)
f(x2)
f(x3)
f
x1
x2
x3
x
g
x
h
x

Hypothesis Space




Explanation of example in terms of Domain Theory:
Stable
Graspable
Liftable
Cup
OpenVessel
BottomIsFlat = T
ConcavityPointsUp = T
Expensive = T
Fragile = T
HandleOnTop = F
HandleOnSide = T
HasConcavity = T
HasHandle = T
Light = T
MadeOfCeramic = T
MadeOfPaper = F
MadeOfStyrofoam = F
0.2
Cup = T
0.8
training derivatives
Target network:
BottomIsFlat
ConcavityPointsUp
Expensive
Fragile
HandleOnTop
HandleOnSide
HasConcavity
HasHandle
Light
MadeOfCeramic
MadeOfPaper
MadeOfStyrofoam
Cup_{target}
Cup

$$E = \sum_i \left[(f(x_i) - \hat{f}(x_i))^2 + \mu_i \sum_j \left(\frac{\partial A(x)}{\partial x^j} - \frac{\partial \hat{f}(x)}{\partial x^j} \right)^2_{(x=x_i)} \right] \tag{1}$$

where

$$\mu_i \equiv 1 - \frac{|A(x_i) - f(x_i)|}{c} \tag{2}$$