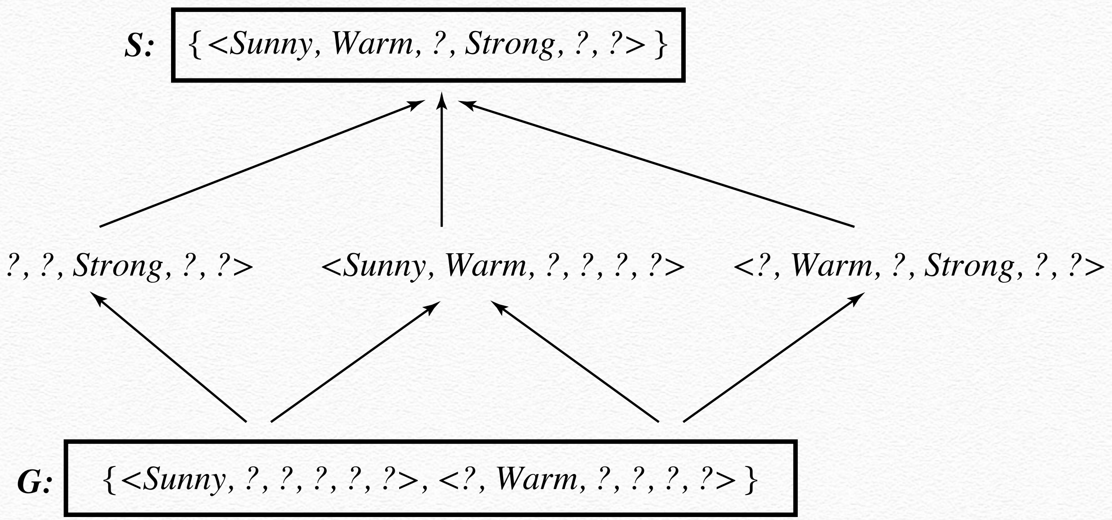
Properties of Candidate-Elimination

- Error/noise in training data (e.g., 2nd training example wrongly labeled as –ve)?
  - Hypotheses inconsistent with 2nd example removed (including target concept c)
  - S and G reduced to Ø with sufficiently large data
- Insufficiently expressive hypothesis representation → biased hypothesis space → $c \notin H$? S and G also reduced to Ø with sufficiently large data

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

Properties of Candidate-Elimination

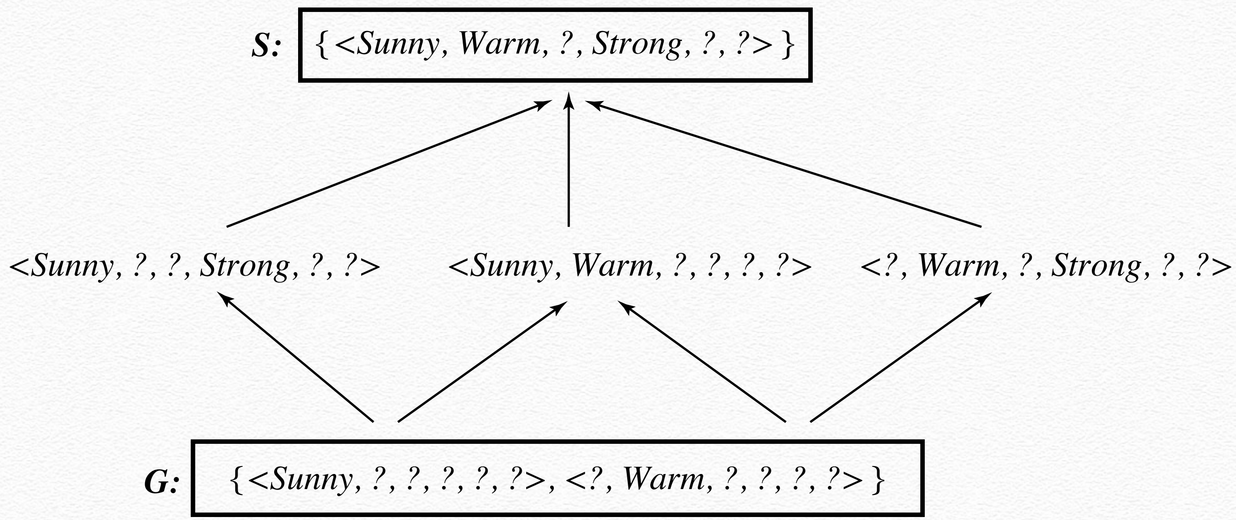
- What input instance should an active learner query next for a training example?
  - Query input instance (e.g., $\langle Sunny, Warm, Normal, Light, Warm, Same \rangle$) that satisfies exactly half of hypotheses in version space (if possible)
  - Version space reduces by half with each training example, hence requiring at least $\lceil \log_2(VS_{H,D}) \rceil$ examples to find target concept c



Properties of Candidate-Elimination

- How to classify new unobserved input instance? What degree of confidence?
  - $\langle Sunny, Warm, Normal, Strong, Cool, Change \rangle$

Proposition 3. An input instance x satisfies every hypothesis in $VS_{H,D}$ iff x satisfies every member of S.



Proof of Proposition 3

$\Leftarrow$ Every input instance x that satisfies every $s \in S$ also satisfies every $h \in VS_{H,D}$.

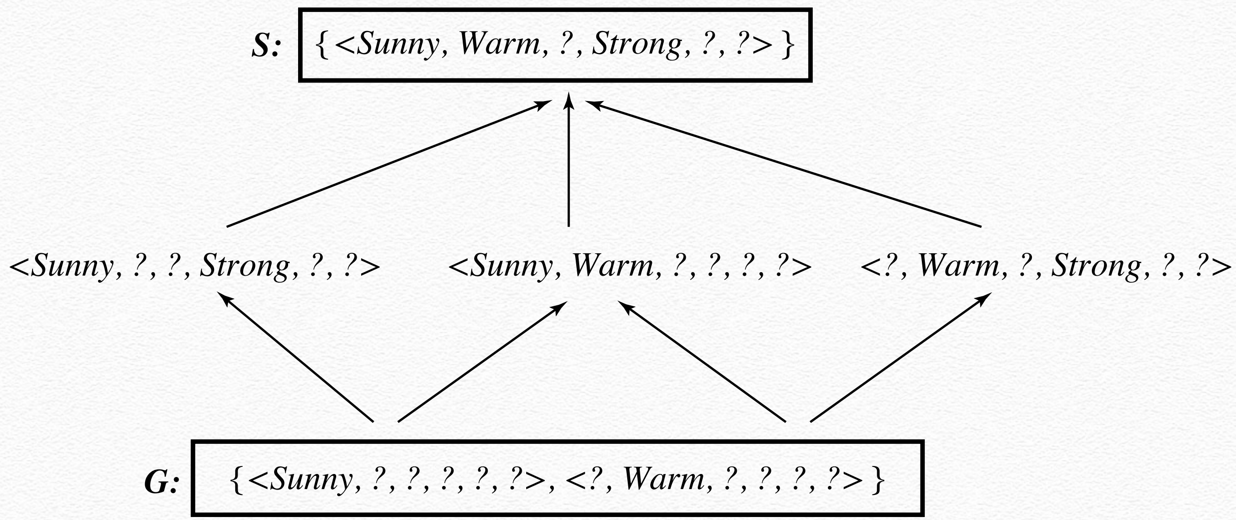
1. $\forall s \in S \;\; s(x) = 1$ is given
2. $\forall h \in VS_{H,D} \; \exists s \in S \;\; h \geq_g s$, by VSRT (page 20)
3. $\forall h \in VS_{H,D} \; \exists s \in S \;\; (s(x) = 1) \rightarrow (h(x) = 1)$, by Def. of $\geq_g$
4. $\forall h \in VS_{H,D} \;\; h(x) = 1$, by steps 1 and 3
5. x satisfies every $h \in VS_{H,D}$

$\Rightarrow$ Every input instance that satisfies every $h \in VS_{H,D}$ also satisfies every $s \in S$. DIY.

Properties of Candidate-Elimination

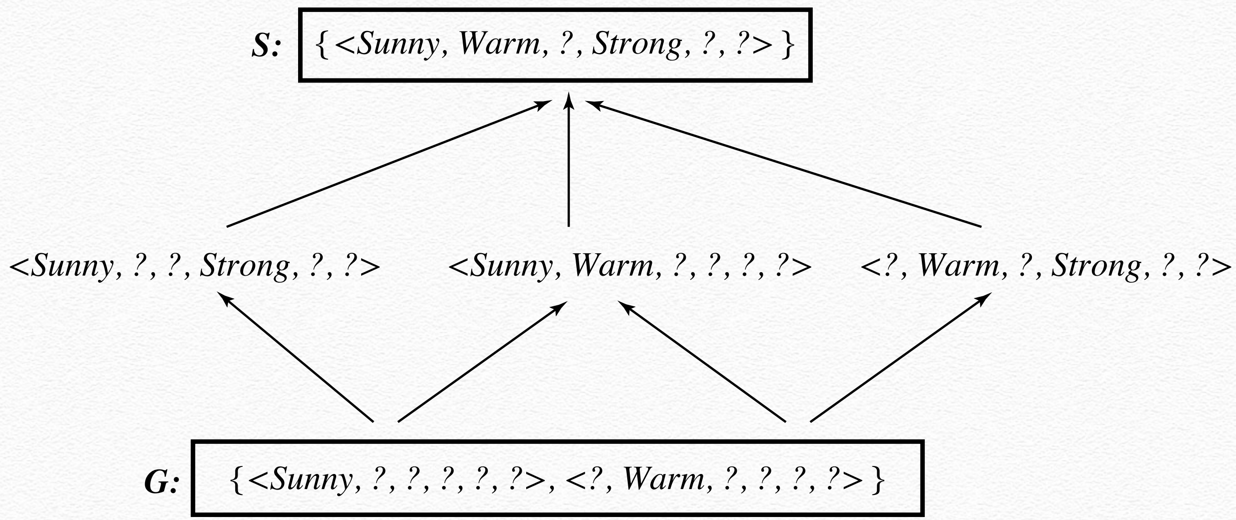
- How to classify new unobserved input instance? What degree of confidence?
  - $\langle \mathit{Rainy}, \mathit{Cool}, \mathit{Normal}, \mathit{Light}, \mathit{Warm}, \mathit{Same} \rangle$

Proposition 4. An input instance x satisfies none of the hypotheses in $VS_{H,D}$ iff x satisfies none of the members of G.



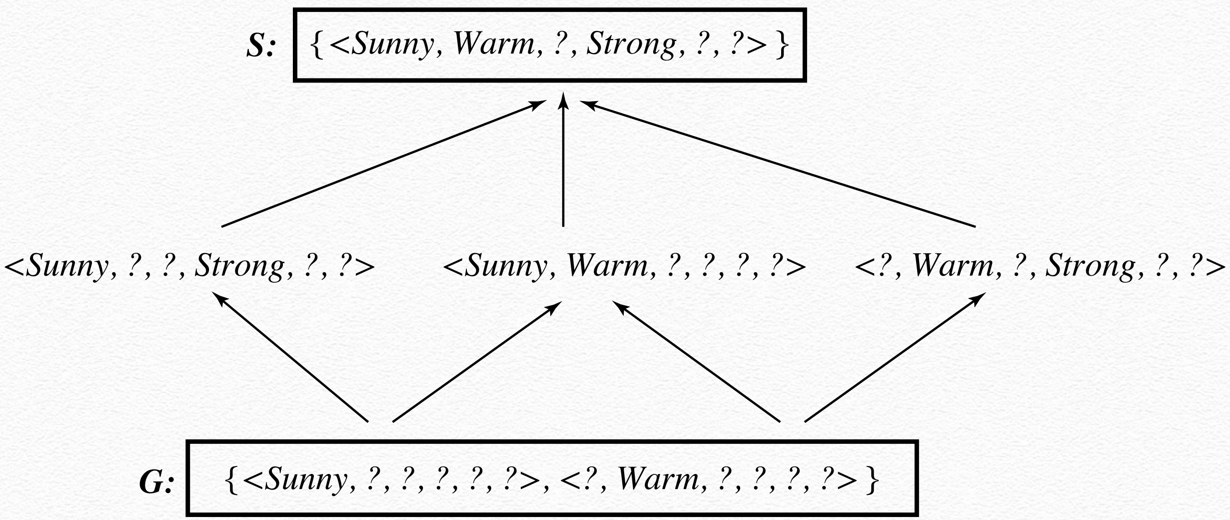
Properties of Candidate-Elimination

- How to classify new unobserved input instance? What degree of confidence?
  - $\langle Sunny, Warm, Normal, Light, Warm, Same \rangle$
  - Optimal query (same input instance as that on page 29)



Properties of Candidate-Elimination

- How to classify new unobserved input instance? What degree of confidence?
  - $\langle Sunny, Cold, Normal, Strong, Warm, Same \rangle$
  - Majority vote is the most probable classification, assuming all hypotheses in *H* are equally probable *a priori*



An Unbiased Learner

Intuition. Choose H that can express every teachable concept (i.e., H is the power set of X)

1. Consider H' = disjunctions, conjunctions, negations of our earlier hypotheses in H for *EnjoySport* task:
   - e.g., $\langle x_1, 1\rangle, \langle x_2, 1\rangle, \langle x_3, 1\rangle, \langle x_4, 0\rangle, \langle x_5, 0\rangle$
   - $S \leftarrow ?$
   - $G \leftarrow ?$
2. Need training examples for every input instance in X to converge to the target concept

Limitation. Cannot classify new unobserved input instances (aka generalize beyond observed training examples)

Inductive Bias

Given

- Concept learning algorithm L
- Input instances X, unknown target concept c
- Noise-free training examples $D_c = \{\langle x_k, c(x_k)\rangle\}_{k=1,\ldots,n}$

Let $L(x, D_c)$ denote the classification of input instance x by L after learning from training examples D_c.

Definition. The **inductive bias** of L is any minimal set of assertions B s.t. for any target concept c and corresponding training examples D_c,

$$\forall x \in X \ \ (B \wedge D_c \wedge x) \vDash (c(x) = L(x, D_c)) \ .$$

Inductive Bias of Candidate-Elimination

Inductive bias of Candidate-Elimination. $B = \{c \in H\}$.

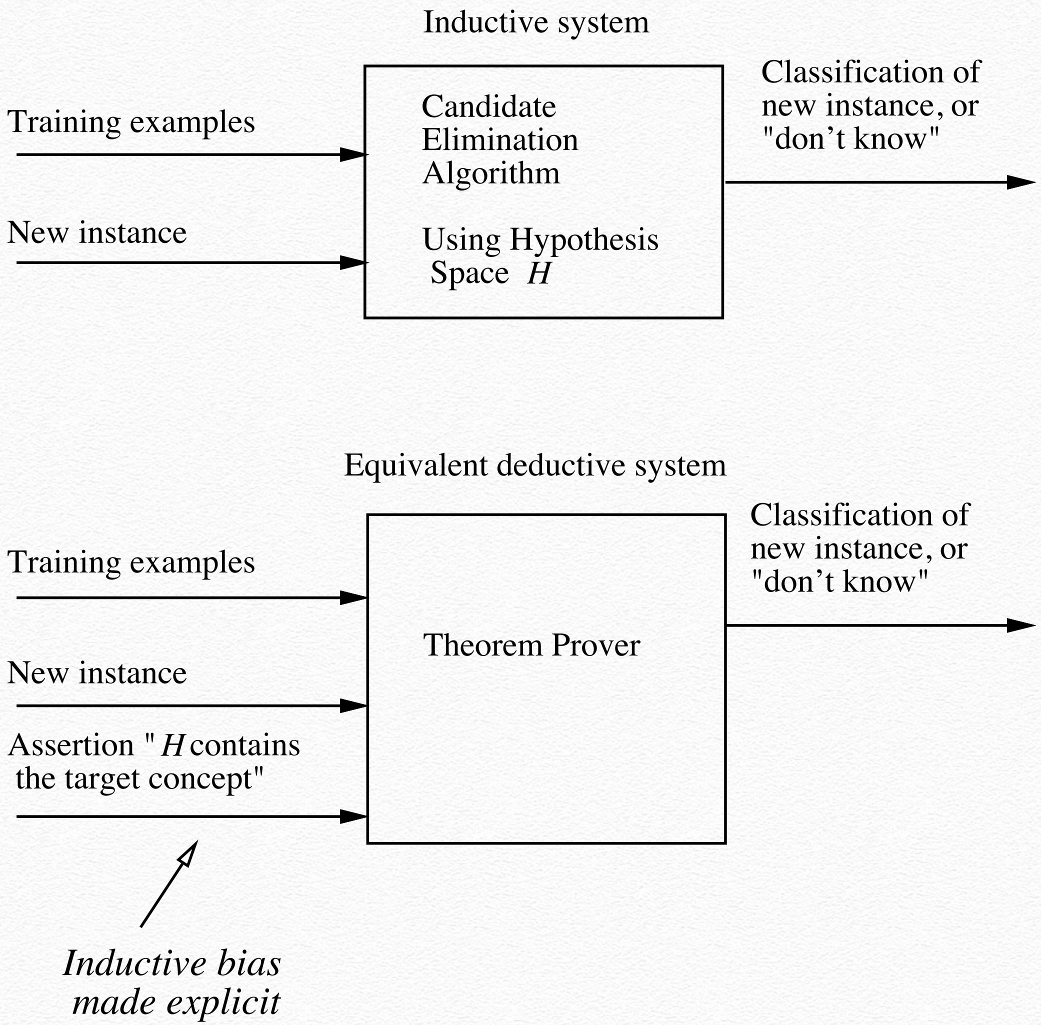
Assumption. Candidate-Elimination outputs a classification $L(x, D_c)$ of input instance x if this vote among hypotheses in VS_{H,D_c} is unanimously +ve or −ve, and does not output a classification otherwise.

1. If $c \in H$, then $c \in VS_{H,D_c}$ since c is consistent with D_c, by Def. of version space (page 17)

2. If L outputs $L(x, D_c)$, then $h(x) = L(x, D_c)$ for every $h \in VS_{H,D_c}$ due to the above assumption, including $c \in VS_{H,D_c}$, by step 1

3. $c(x) = L(x, D_c)$

Inductive vs. Deductive Inference



Comparing Learners with Different Inductive Biases

- Rote-Learner. Store examples & classify input instance x iff it matches that of previously observed example. No inductive bias

- Candidate-Elimination. Inductive bias: $c \in H$

- Find-S. Inductive bias: $c \in H$ and all instances are –ve unless the opposite is entailed by its other knowledge

Summary

- Concept learning as search through *H*
- General-to-specific ordering over *H*
- Candidate elimination algorithm
- Boundaries *S* and *G* characterize learner's uncertainty
- Active learner can generate informative queries
- Stronger inductive bias allows classification of greater proportion of unobserved input instances
- Inductive learner can be modeled by an equivalent deductive inference system