



# Concept Learning

Aliridho Barakbah, Entin Martiana

Knowledge Engineering Research Group  
Department of Information and Computer Engineering  
Politeknik Elektronika Negeri Surabaya



# Ide Mesin Pembelajaran

Fakta harian dalam 6 hari dan keputusan untuk berolah-raga sebagai berikut:

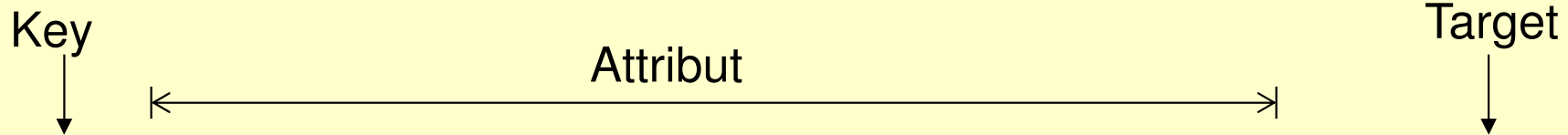
#	Cuaca	Temperatur	Kecepatan Angin	Berolah-raga
1	Cerah	Normal	Pelan	Ya
2	Cerah	Normal	Pelan	Ya
3	Hujan	Tinggi	Pelan	Tidak
4	Cerah	Normal	Kencang	Ya
5	Hujan	Tinggi	Kencang	Tidak
6	Cerah	Normal	Pelan	Ya

(1) *Ketika cuaca cerah, apakah akan berolah-raga?*

(2) *Ketika cuaca cerah dan temperatur normal, apakah akan berolah-raga?*

Penyajian keputusan berdasarkan fakta inilah yang mengilhami konsep dari mesin pembelajaran

# Data Training



Day	Cuaca	Temperatur	Kecepatan Angin	Berolah-raga
D1	Cerah	Normal	Pelan	Ya
D2	Cerah	Normal	Pelan	Ya
D3	Hujan	Tinggi	Pelan	Tidak
D4	Cerah	Normal	Kencang	Ya
D5	Hujan	Tinggi	Kencang	Tidak
D6	Cerah	Normal	Pelan	Ya

- Attribut adalah kolom data, ada atribut dan target
- Instance adalah isi dari atribut sebagai contoh atribut cuaca mempunyai instance “cerah” dan “hujan”, sering ditulis dengan cuaca={cerah,hujan}
- Record/tuple adalah baris data



# Ide Mesin Pembelajaran

---

Pada dasarnya semua algoritma yang dikembangkan dalam mesin pembelajaran adalah algoritma yang menghasilkan hipotesa dari suatu keputusan berdasarkan data pembelajaran yang diberikan.



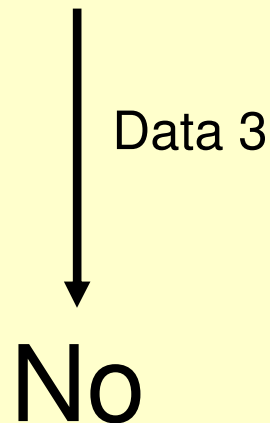
# Fact

Data	Sky	AirTemp	Humidity	Wind	Water	Forecast	EnjoySport
1	Sunny	Warm	Normal	Strong	Warm	Same	Yes
2	Sunny	Warm	High	Strong	Warm	Same	Yes
3	Rainy	Cold	High	Strong	Warm	Change	No
4	Sunny	Warm	High	Strong	Cool	Change	Yes

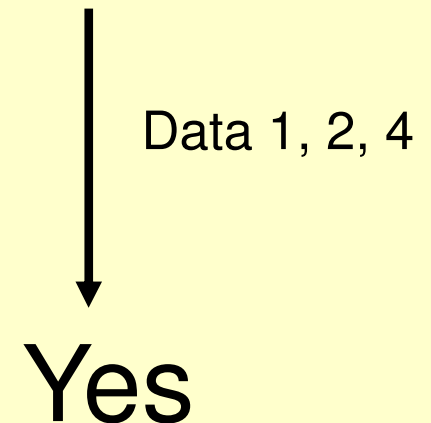
# Problem description

---

<?, Cold, High, ?, ?, ?>



<Sunny, Warm, ?, ?, ?, ?>



**Our human brain can answer these questions.  
But how the machine can answer?**



# Contoh Keputusan Dari Hipotesa

Bila cuaca cerah,  
apakah akan  
berolahraga?



## HIPOTESA

$H(\text{Sunny, ?, ?, ?, ?, ?}) = \text{Ya}$

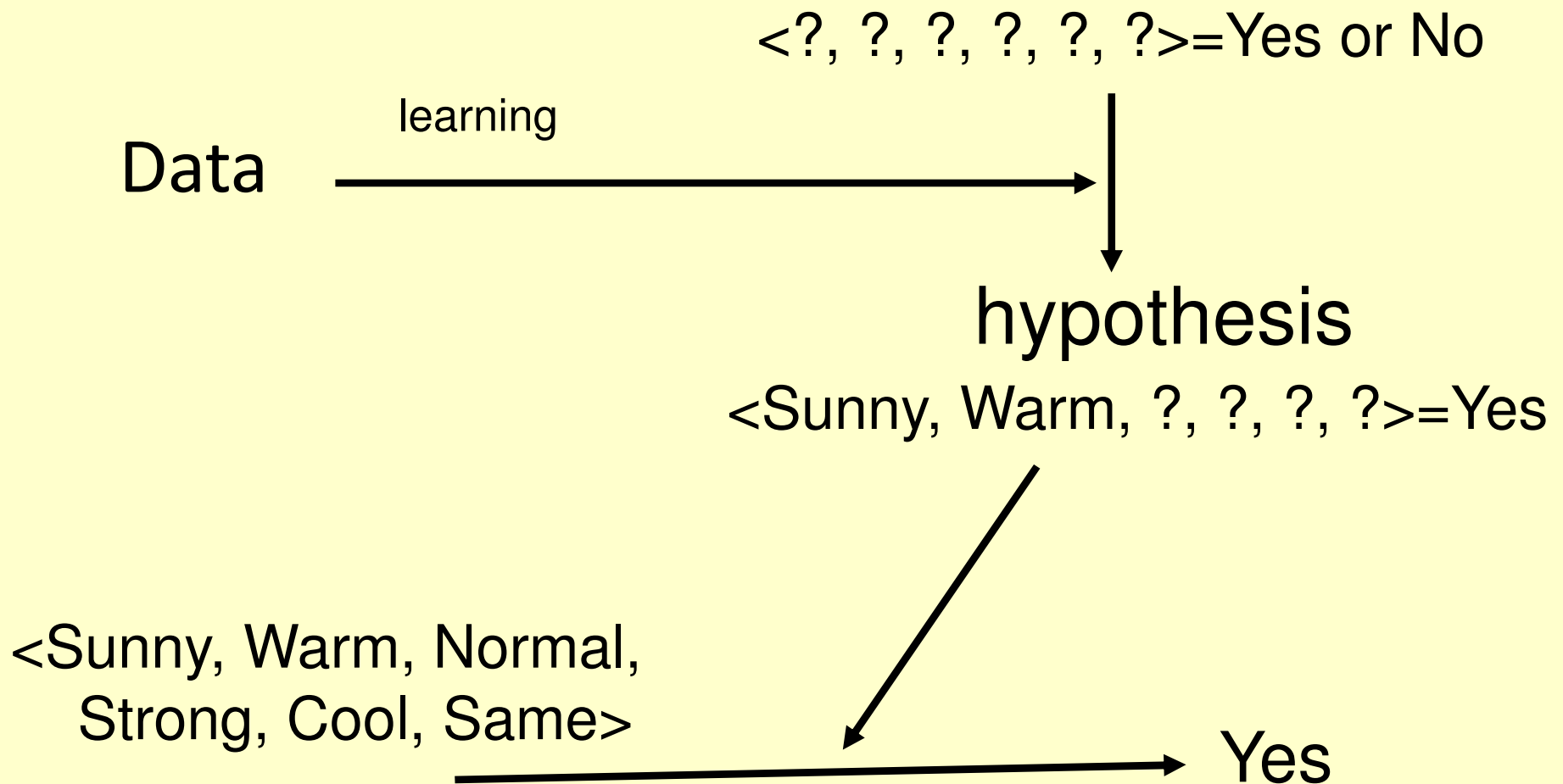
$H(\text{?, Cold, High, ?, ?, ?}) = \text{Tidak}$

CEK

**JAWAB : YA**



# Learning Process





# Find-S

---

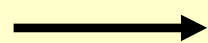
- Find-S adalah suatu metode paling sederhana yang dapat digunakan untuk mendapatkan suatu hipotesa berdasarkan data.
- Find-S mencari kesamaan nilai atribut untuk memperoleh suatu hipotesa
- Kelemahan dari Find-S adalah data yang digunakan harus bersifat konsisten dan tidak bias ??? (Terlalu sulit untuk dapat memperoleh data semacam ini pada persoalan nyata)



# Find-S

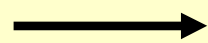
<  $\phi$ ,  $\phi$ ,  $\phi$ ,  $\phi$ ,  $\phi$ ,  $\phi$  >

<Sunny, Warm, Normal,  
Strong, Warm, Same>



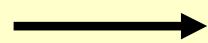
<Sunny, Warm, Normal, Strong, Warm, Same>

<Sunny, Warm, High,  
Strong, Warm, Same>



<Sunny, Warm, ? , Strong, Warm, Same>

<Sunny, Warm, High,  
Strong, Cool, Change>



<Sunny, Warm, ? , Strong, ? , ? >



- 
- Advantage
    - Very simple
  - Disadvantage
    - Ignores the negative data



# Candidate-Elimination

---

$S_0$   $\langle \phi, \phi, \phi, \phi, \phi, \phi \rangle$

?

$G_0$   $\langle ?, ?, ?, ?, ?, ? \rangle$



# Candidate-Elimination

$S_0$   $\langle \phi, \phi, \phi, \phi, \phi, \phi \rangle$

$S_1$   $\langle \text{Sunny, Warm, Normal, Strong, Warm, Same} \rangle$

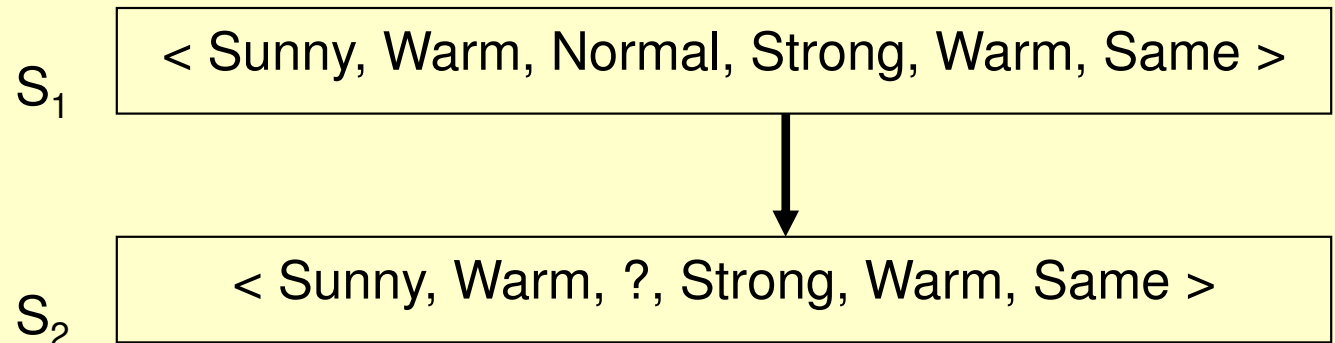
$\langle \text{Sunny, Warm, Normal, Strong, Warm, Same} \rangle$   
=Yes

?

$G_1$   $\langle ?, ?, ?, ?, ?, ? \rangle$

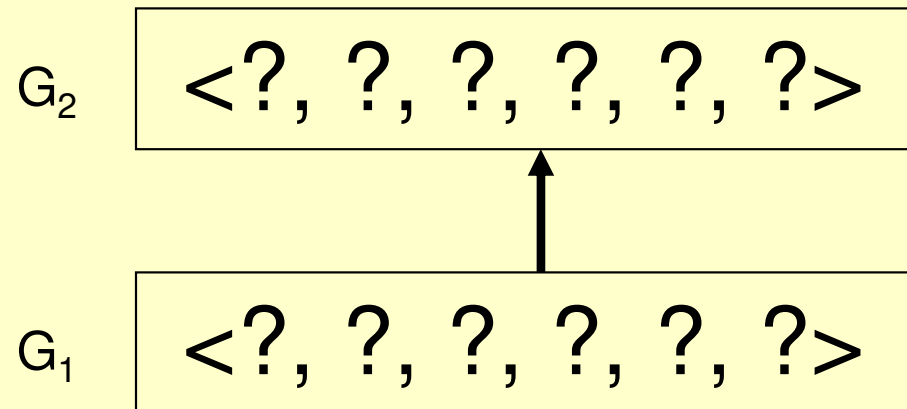
$G_0$   $\langle ?, ?, ?, ?, ?, ? \rangle$

# Candidate-Elimination



?

<Sunny, Warm, High,  
Strong, Warm, Same>  
=Yes



# Candidate-Elimination

<Rainy, Cold, High,  
Strong, Warm, Change>  
=No

$S_2$  < Sunny, Warm, ?, Strong, Warm, Same >



$S_3$  < Sunny, Warm, ?, Strong, Warm, Same >

?

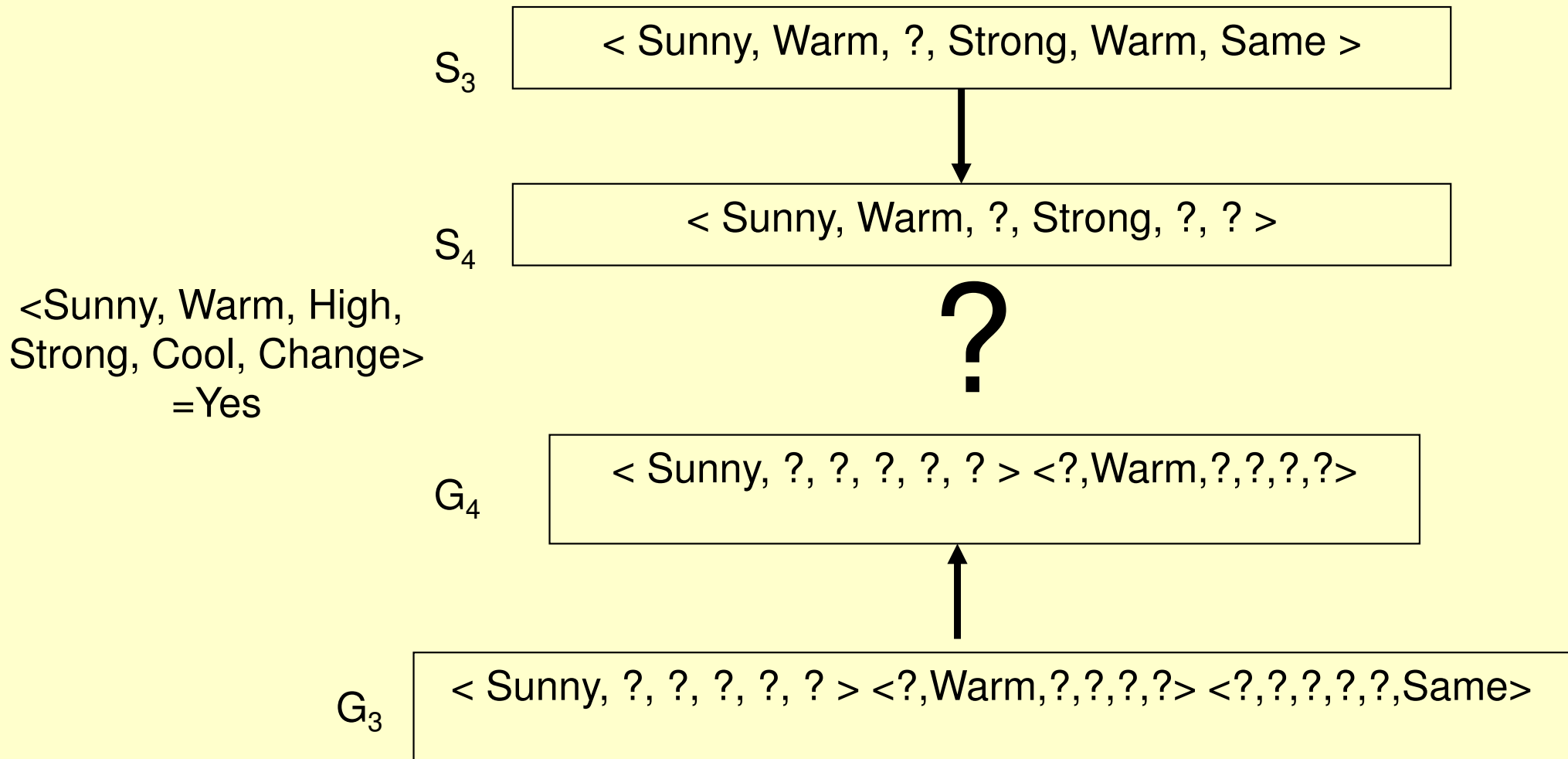
$G_3$  < Sunny, ?, ?, ?, ?, ? > < ?, Warm, ?, ?, ?, ? > < ?, ?, ?, ?, ?, Same >



$G_2$  < ?, ?, ?, ?, ?, ? >

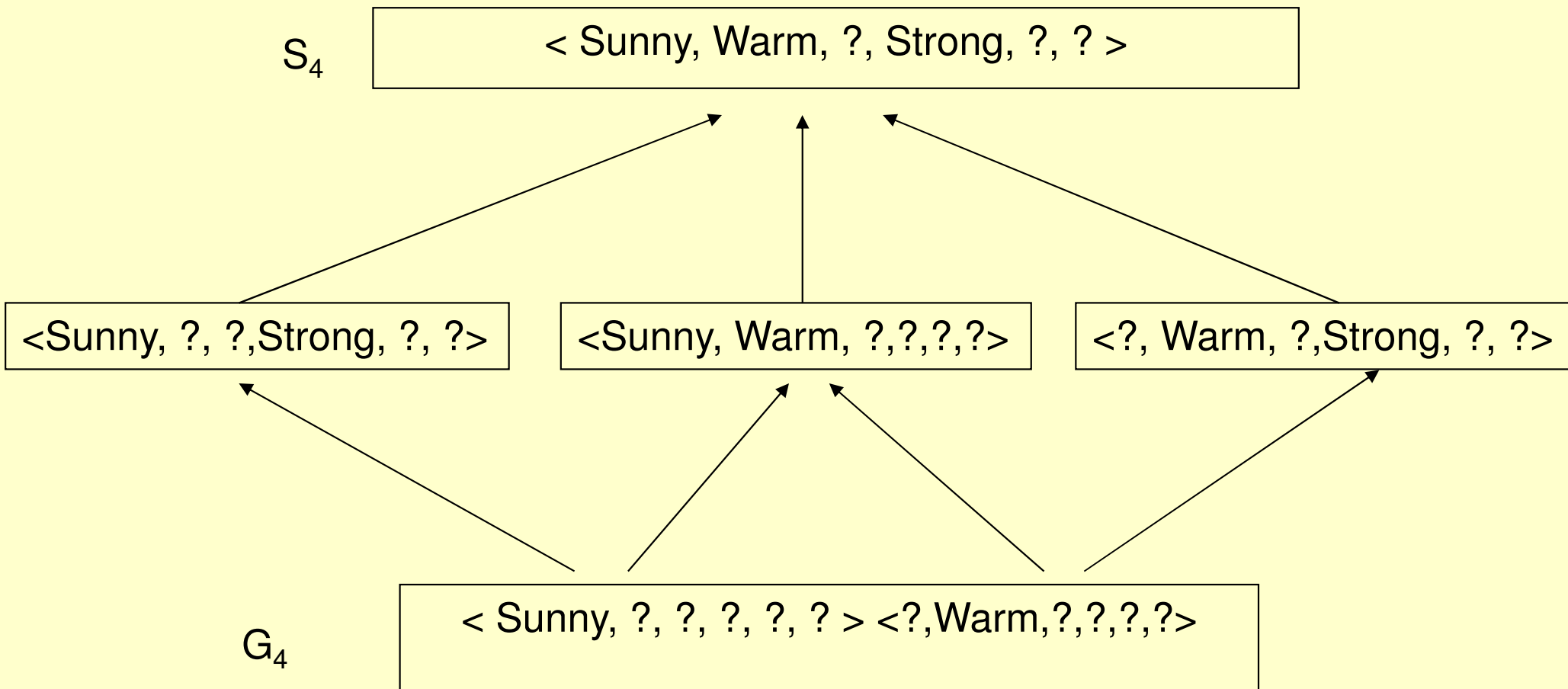


# Candidate-Elimination





# Candidate-Elimination



- 
- Advantage
    - Consider the negative data to strengthen the hypothesis
  - Disadvantage
    - If the data is not consistent, S and G can not match
    - Difficult to implement in the programming



# Latihan - Pengenalan buah

Panjang	Lebar	Nama buah
panjang	sedang	Pisang
panjang	pendek	Pisang
pendek	pendek	Apel

Buatlah hipotesis untuk Pisang dan Apel dengan melakukan Concept Learning menggunakan Find-S

# Latihan - Penyakit hipertensi

Umur	Kegemukan	Hipertensi
muda	gemuk	Tidak
muda	sangat gemuk	Tidak
paruh baya	gemuk	Tidak
paruh baya	terlalu gemuk	Ya
tua	terlalu gemuk	Ya

Buatlah hipotesis untuk Tidak dan Ya pada kasus penyakit Hipertensi dengan melakukan Concept Learning menggunakan Find-S

