Concept Learning

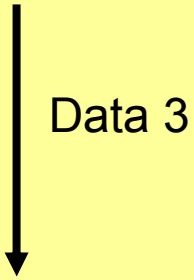
Ali Ridho Barakbah

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

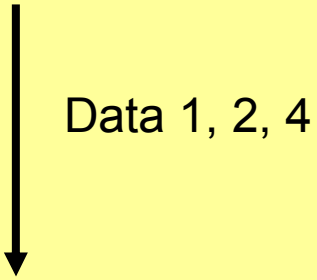
Problem description

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



No

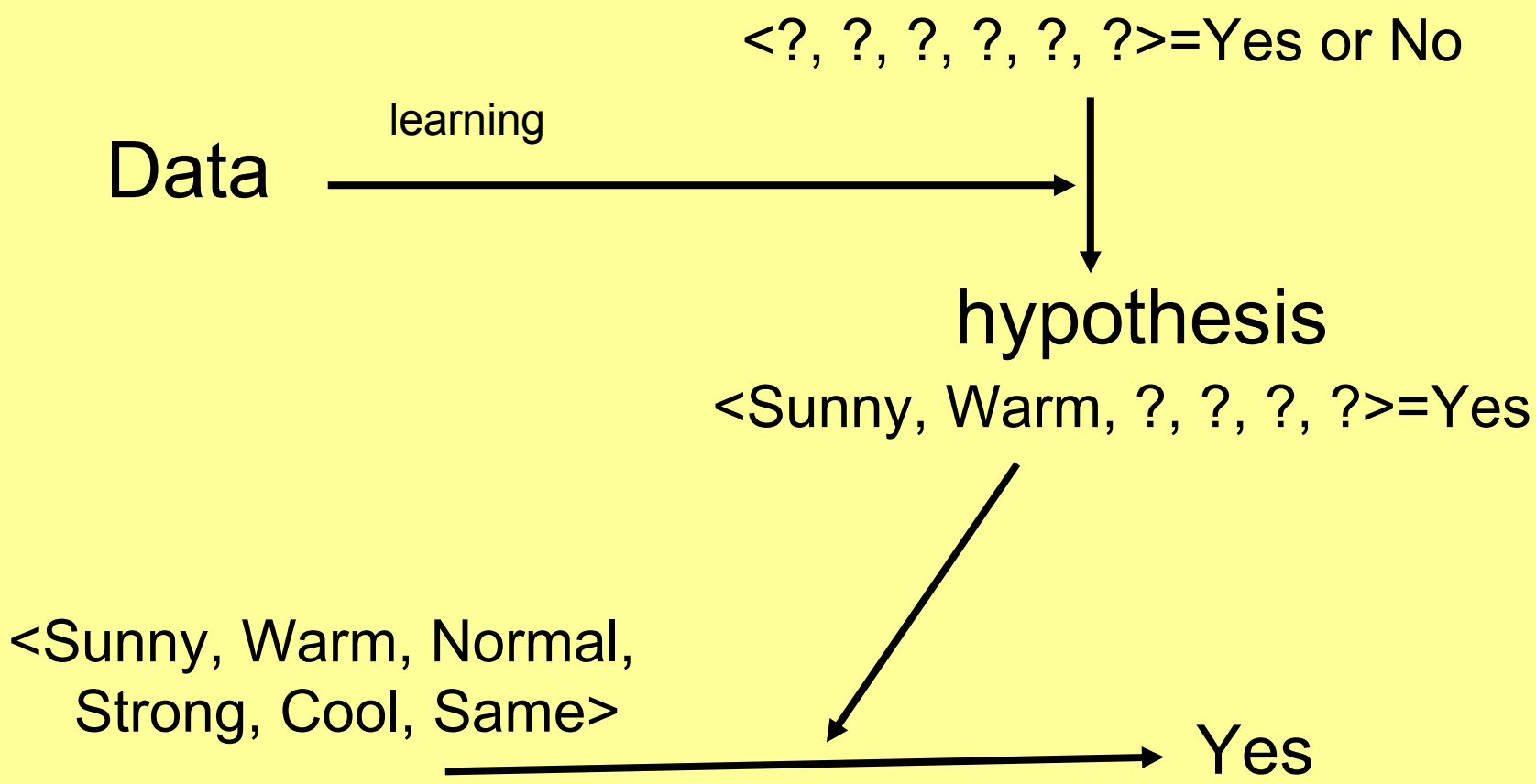
<Sunny, Warm, ?, ?, ?, ?>



Yes

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

Learning Process



Find-S

< ϕ, ϕ, ϕ, ϕ, ϕ, ϕ >

<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

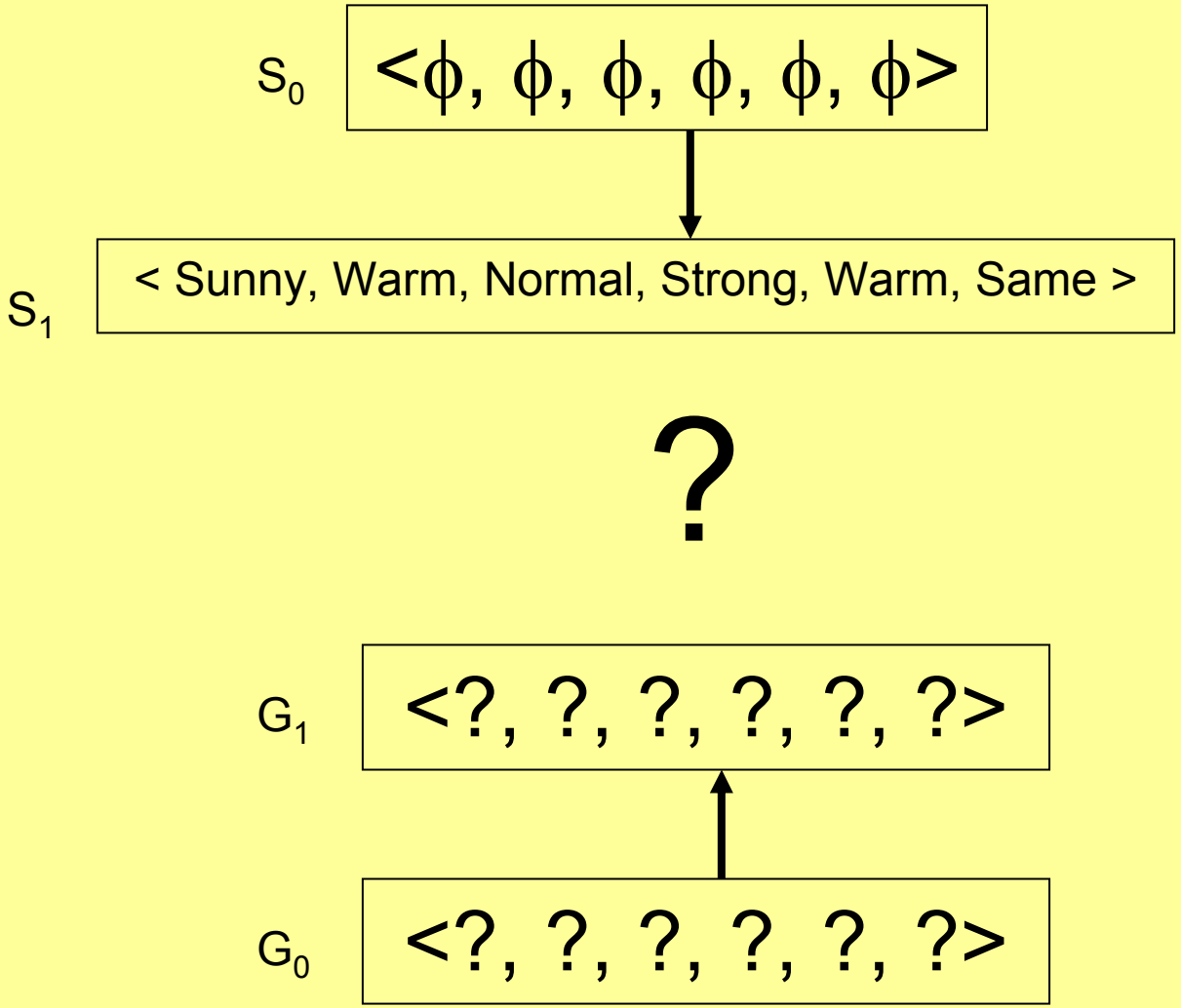
Candidate-Elimination

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

?

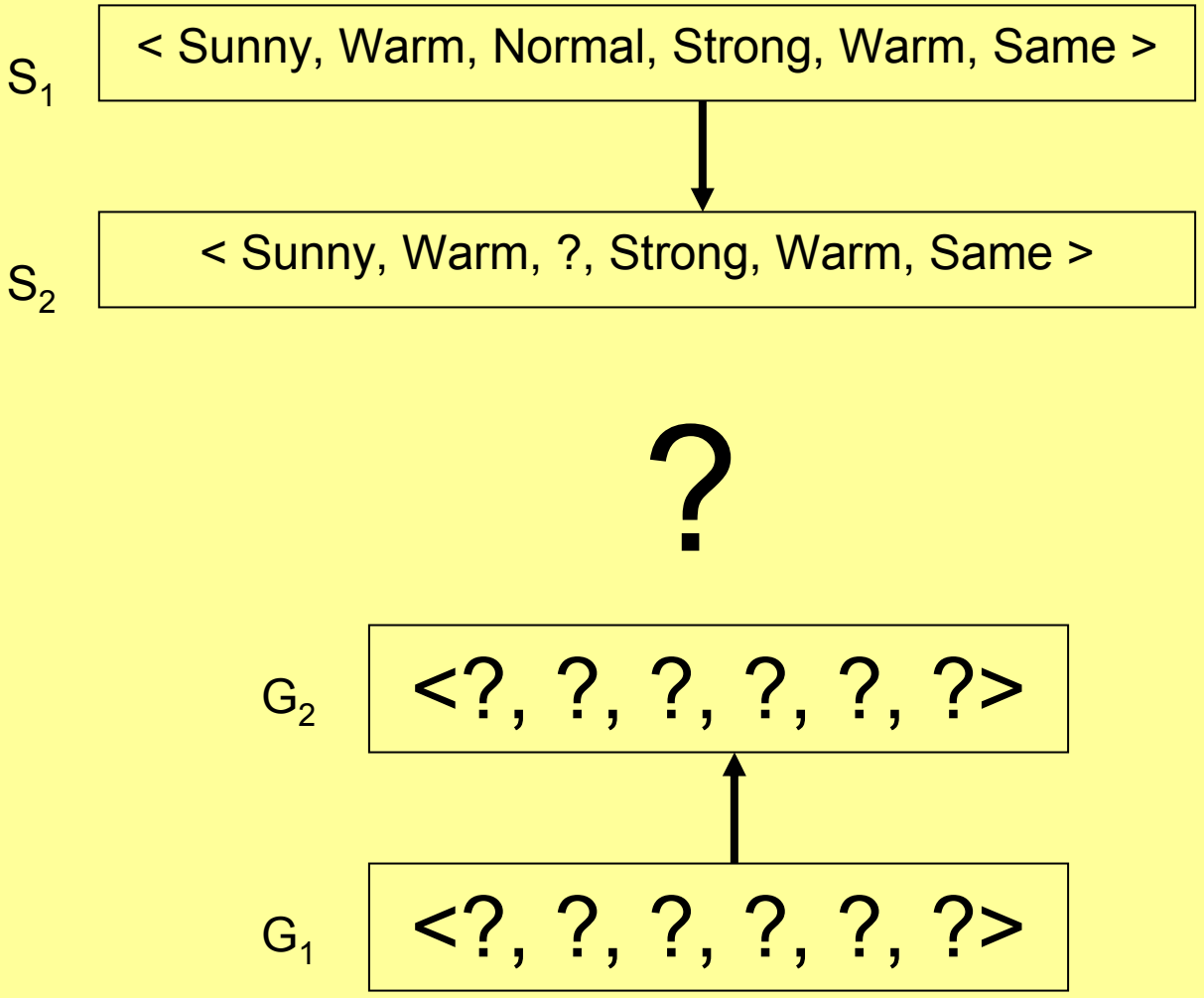
G_0 $\langle ?, ?, ?, ?, ?, ? \rangle$

Candidate-Elimination

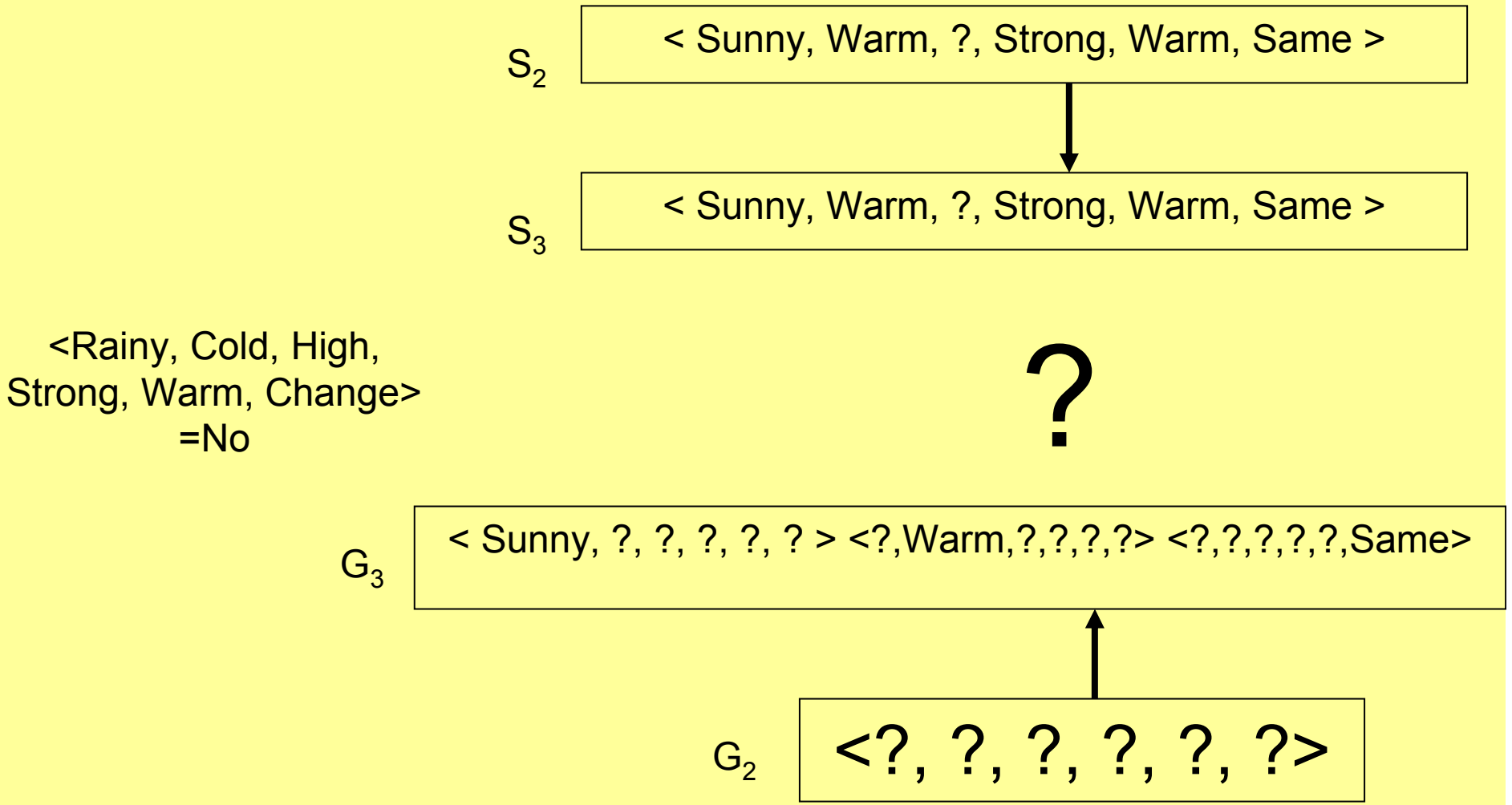


<Sunny, Warm, Normal,
Strong, Warm, Same>
=Yes

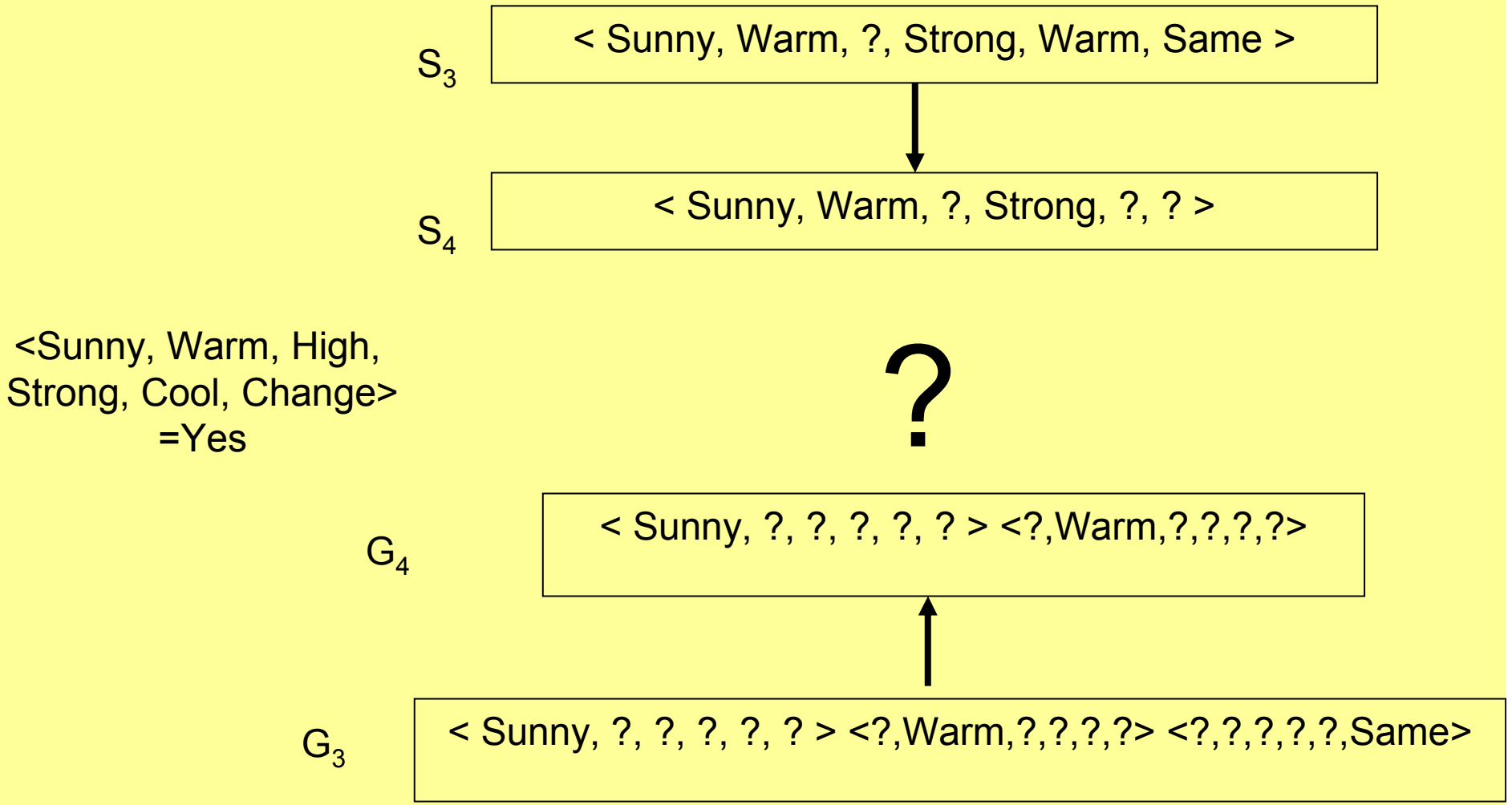
Candidate-Elimination



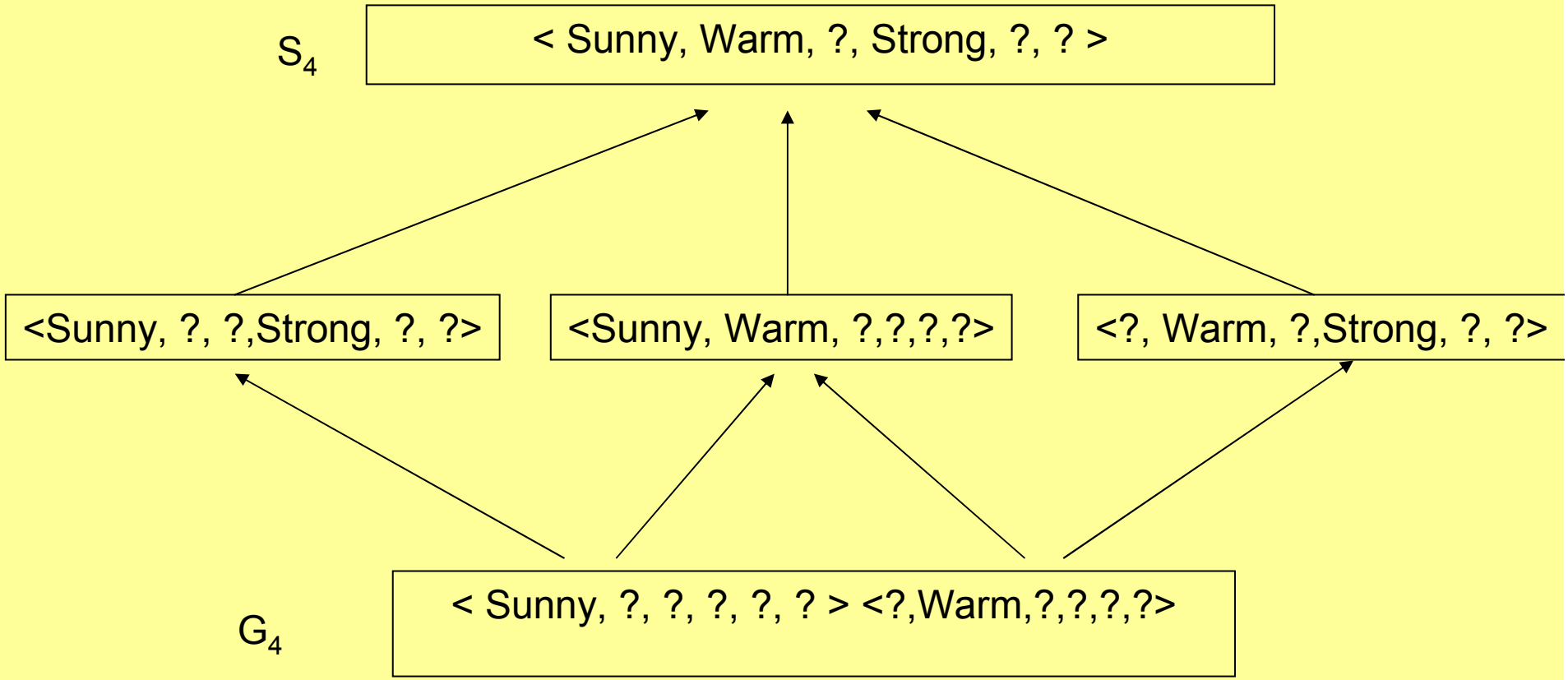
Candidate-Elimination



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