# 2-2. introduction to Rule Based Algorithm

# Find s 알고리즘

# 

# Xi 나가 놀은 instance라 했을 때

# x1: <Sunny, Warm, Normal, Strong, Warm, Same>

# x2: <Sunny, Warm, Normal, Light, Warm, Same>

# • x4: <Sunny, Warm, Normal, Strong, Warm, Change>

# • Hypotheses

# • h0=<∅,∅,∅,∅,∅,∅>

# • h1=<Sunny, Warm, Normal, Strong, Warm, Same> • h1,2,3=<Sunny, Warm, Normal, ?, Warm, Same>

# • h1,2,3,4=<Sunny, Warm, Normal, ?, Warm, ?>

# • Any problems?

# • Many possible hs, and can’t determine the converge

# 정확하게 있는것만 찝어내는 방식이다.

# 허나 다양한 데이터가 존재할시 가설이 정말 다양하게 나올수 밖에 없다.

# Version Space 일어날 가능성 범위 지정

# 다양한 가설을 모아둔 것 어떤하나의 가설로 모아지기는 불가능

# 이정도의 가설은 가능하지 않나 범위를 지정

# General한 범위 <<<<<<<<<<<<<<<<<<

# • Many hypotheses possible, and No way to find the convergence • Need to setup the perimeter of the possible hypothesis • The set of the possible hypotheses == Version Space, VS • General Boundary, G • Is the set of the maximally general hypotheses of the version space • Specific Boundary, S • Is the set of the maximally specific hypotheses of the version space • Every hypothesis, h, satisifies

# G: {<Sunny, ?,?,?,?,?>, <?,Warm, ?, ?, ?, ?>} 이게 versionspace가 됩니다.

# Candidate Elimination Algorithm

# 가장 구체적인 가설과 가장 보편적인 가설을 추려내어서 이를 고려해서 versionspace

# Initialize S to maximally specific h in H

# • Initialize G to maximally general h in H

# • For instance x in D

# • If y of x is positive

# • Generalize S as much as needed to cover o in x

# S0: {<∅,∅,∅,∅,∅,∅>} 가장 구체적인 예

# • Remove any h in G, for which h(o)≠y

# G0: {<?,?,?,?,?,?>} 가장 보편적인 예

# • If y of x is negative

# • Specialize G as much as needed to exclude o in x

# • Remove any h in S, for which h(o)=y

# 예시

# S0: {<∅,∅,∅,∅,∅,∅>}

# S1: {<Sunny, Warm, Normal, Strong, Warm, Same>}

# S2: {<Sunny, Warm, ?, Strong, Warm, Same>}

# S4: {<Sunny, Warm, ?, Strong, ?, ?>}

# G4: {<Sunny,?,?,?,?,?>, <?,Warm,?,?,?,?>}

# G3: {<Sunny,?,?,?,?,?>, <?,Warm,?,?,?,?>, <?,?,?,?,?,Same>

# G0, G1, G2: {<?,?,?,?,?,?>}

# 

# 

# 최종적으로 나온 결과인 s와 g에 새 데이터 3개를 넣어보자

<Sunny, Warm, Normal, Strong, Cool, Change> 0

• <Rainy, Cold, Normal, Light, Warm, Same> 0

• <Sunny, Warm, Normal, Light, Warm, Same> x

3번째 g는 맞지만 s는 틀린다.

1. Is this working?

결국 과거의 데이터분석 학습 기술로는 위의 3번째 데이터 오류처럼 완벽하지가 않다.