Homework: Version Spaces

Seppe Duwé
r0349304

November 19, 2015

Exercise 1. A single man is searching for a girlfriend and creates a profile on the online dating platform "Tinder". Instead of judging the picture, he has a conversation with the first five candidates, of which two conversations where nice. He wonders whether he can automate the selection process to save unnecessary work by looking at there interest, age, eye color, picture and relation status.

The examples are:

Interest	Age	Eye Color	Picture	Relation	Like
Stone Collection	17	Blue	Handsome	Bachelor	No
Traveling	23	Brown	Neutral	Bachelor	Yes
Cycling	31	Amber	Handsome	Relation	No
Reading	26	Blue	Handsome	Bachelor	Yes
Reading	25	Green	Ugly	Relation	No

The interests, ages and eye colors are organised in general-to-specic hierarchies, given in the figures below. This means that e.g. the hypothesis [competition, young, ?, ?, ?] is more general than [competition, 19, mono, ?, ?], which in its turn is more general than [chess, 19, mono, handsome,?]. Apply the Version Space algorithm on this example. Indicate per example how the spaces G and S evolve. Clearly indicate which hypotheses are pruned away, and why.

What is (or what are) the concept(s) that are deducted ? Using this result, what conclusion can he make for the following additional profiles (should he talk to the girl or not, or is it undecided ?)

[Stone Collection,31,Green,Handsome,Relation]
[Traveling,19,Blue,Ugly,Bachelor]
[Chess,25,Amber,Neutral,Bachelor]

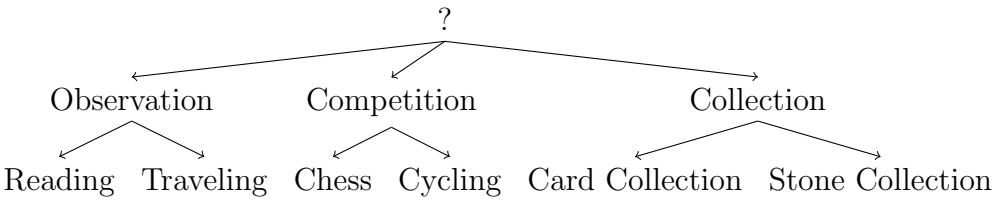


Figure 1: Interest hierarchy for Example 1.

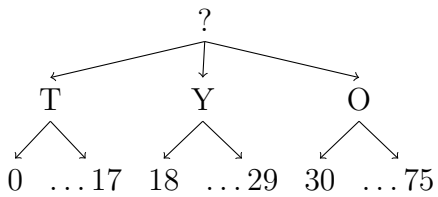


Figure 2: Age hierarchy (TY= too young, Y=young, O= old) for Example 1.

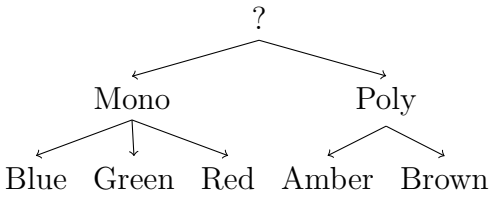


Figure 3: Eye color hierarchy for Example 1.

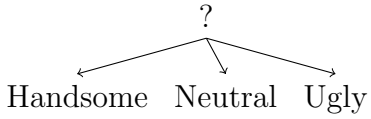


Figure 4: Picture hierarchy for Example 1.

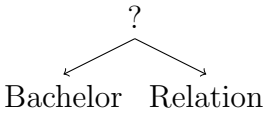
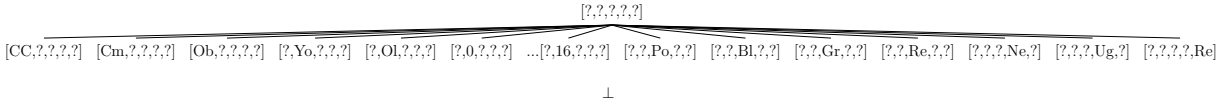
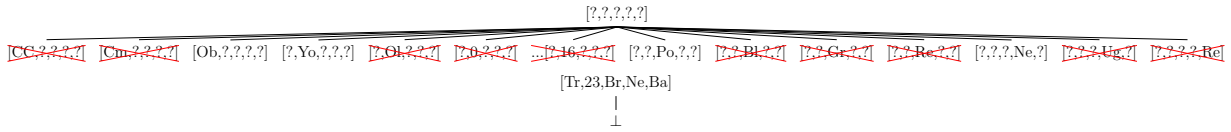


Figure 5: Relation hierarchy for Example 1.

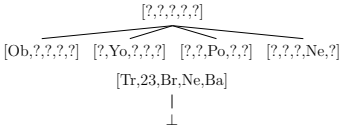
Solution Exercise 1. [Stone Collection,17,Blue,Handsome,Bachelor] -



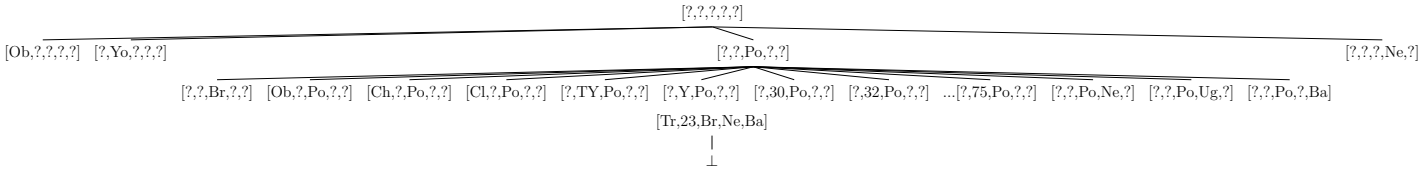
[Traveling,23,Brown,Neutral,Bachelor] +



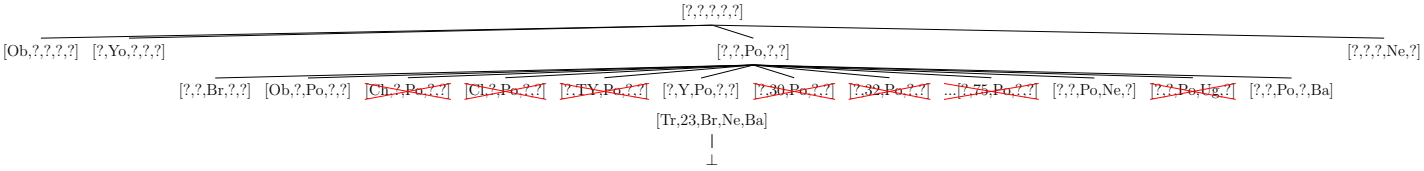
24 out of 28 do not cover last positive example



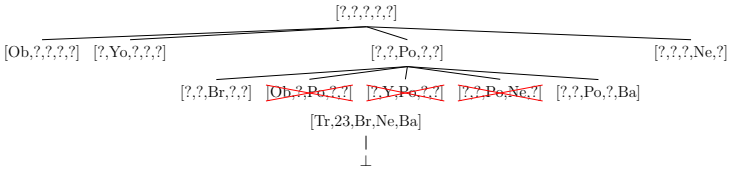
[Cycling,31,Amber,Handsome,Relation] -



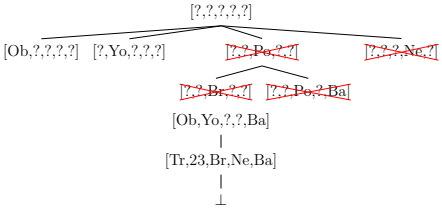
49 out of 54 do not generalize the specific model



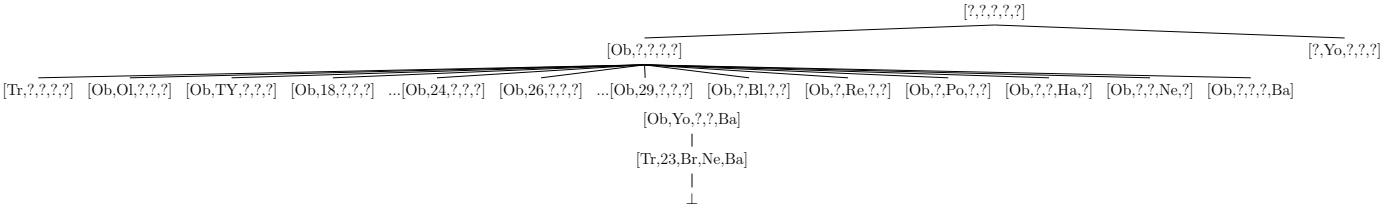
Redundant hypothese



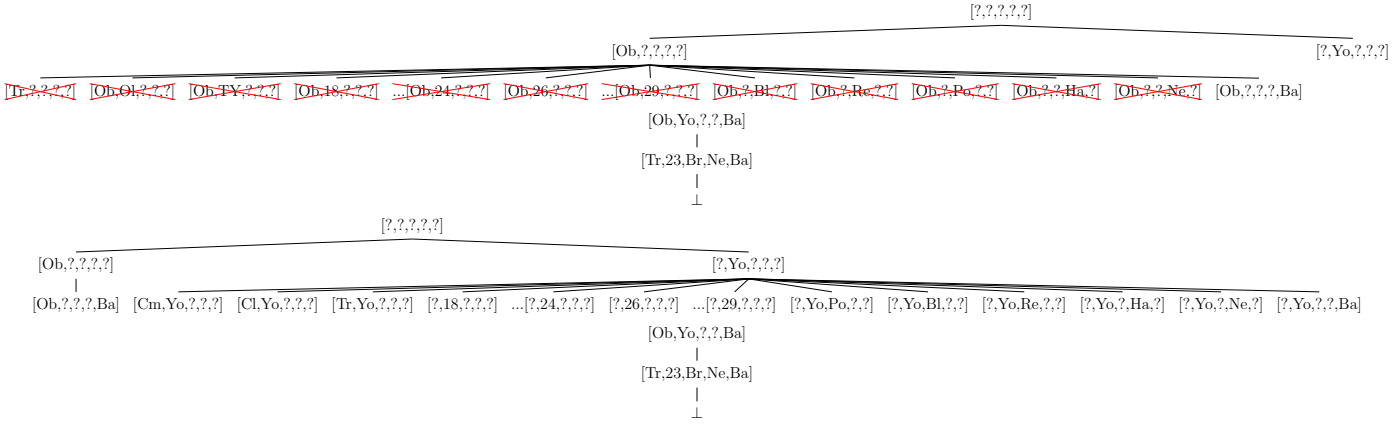
[Reading,26,Blue,Handsome,Bachelor] +
4 out of 6 do not cover last positive example



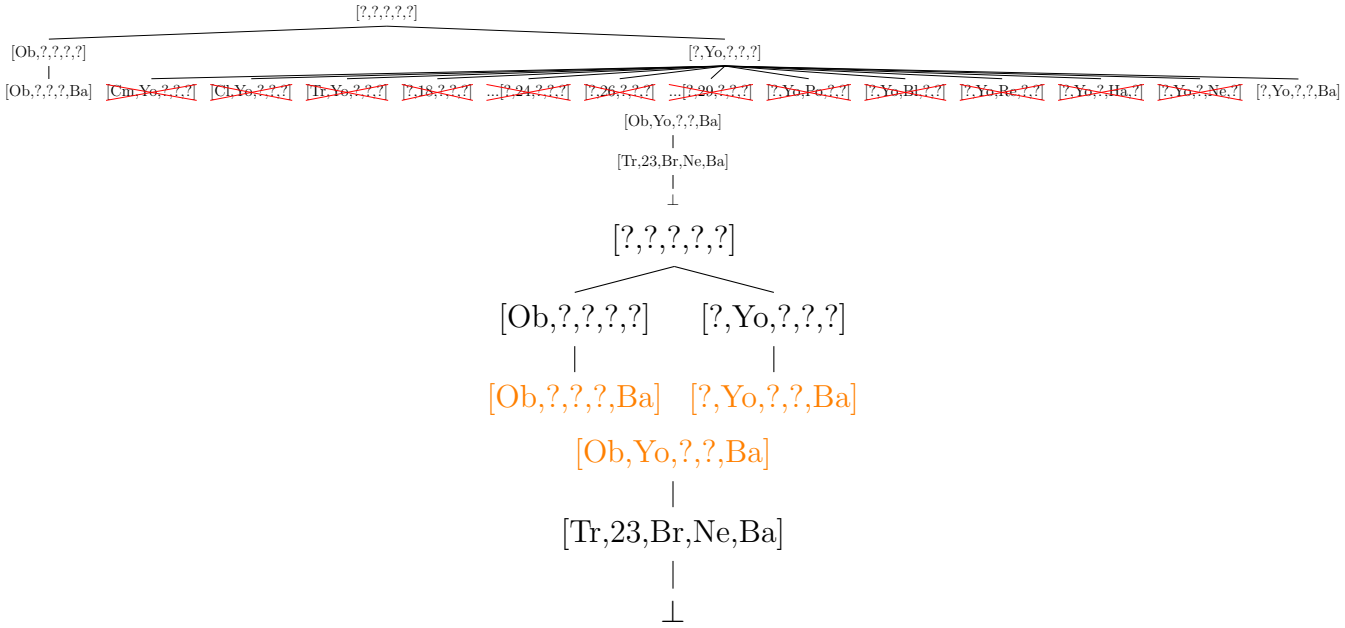
[Reading,25,Green,Ugly,Relation] -



19 out of 20 do not generalize the specific model



19 out of 20 do not generalize the specific model



[Stone Collection,31,Green,Handsome,Bachelor]:No

- Not More Specific than [Ob,?,?,?,Ba]
- Not More Specific than [?,Yo,?,?,Ba]

[Traveling,19,Blue,Ugly,Bachelor]:Yes

- More Specific than [Ob,?,?,?,Ba]
- More Specific than [Ob,Yo,?,?,Ba]

[Chess,25,Amber,Neutral,Bachelor]:Maybe

- More Specific than [?,Yo,?,?,Ba]
- Not More Specific than [Ob,Yo,?,?,Ba]