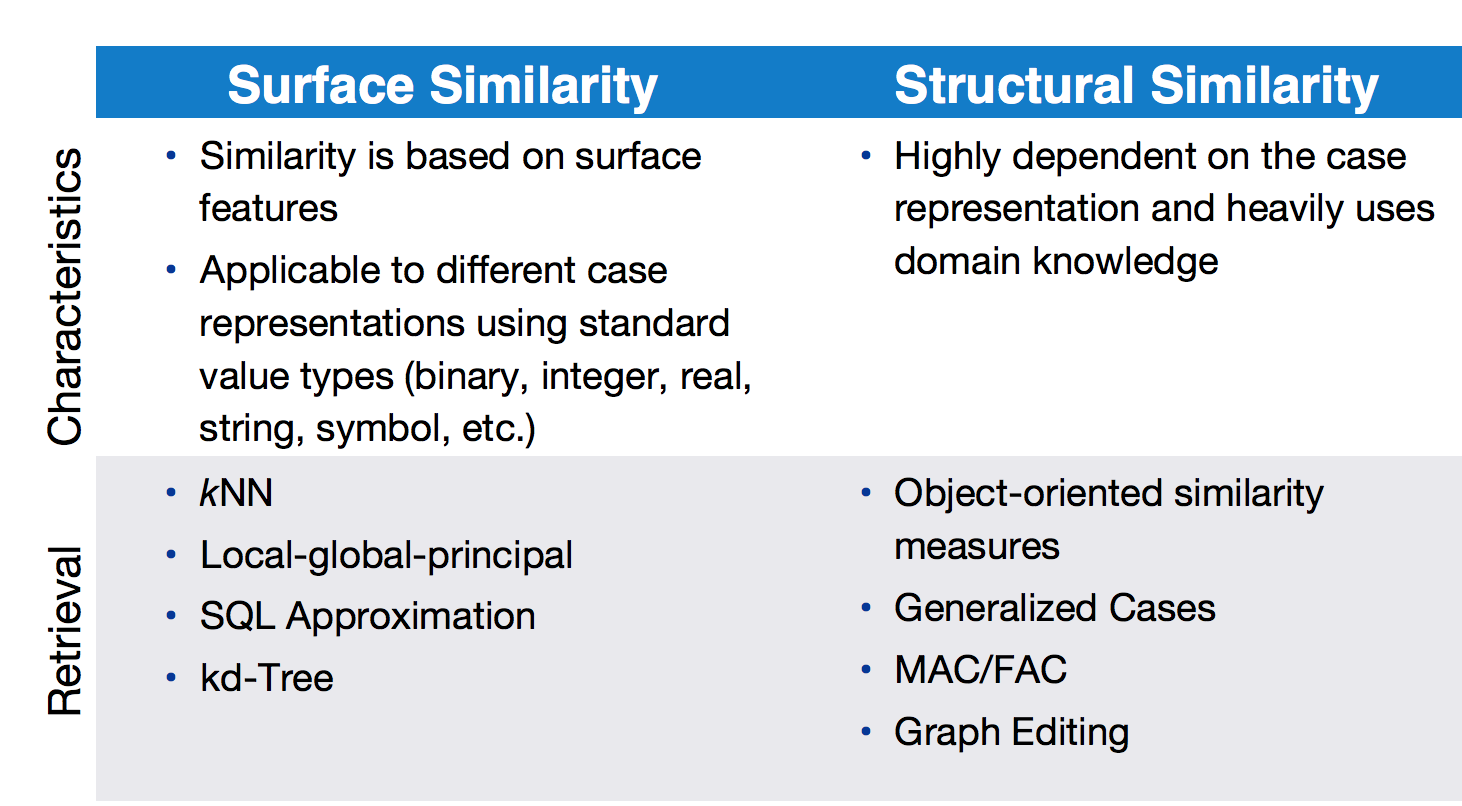
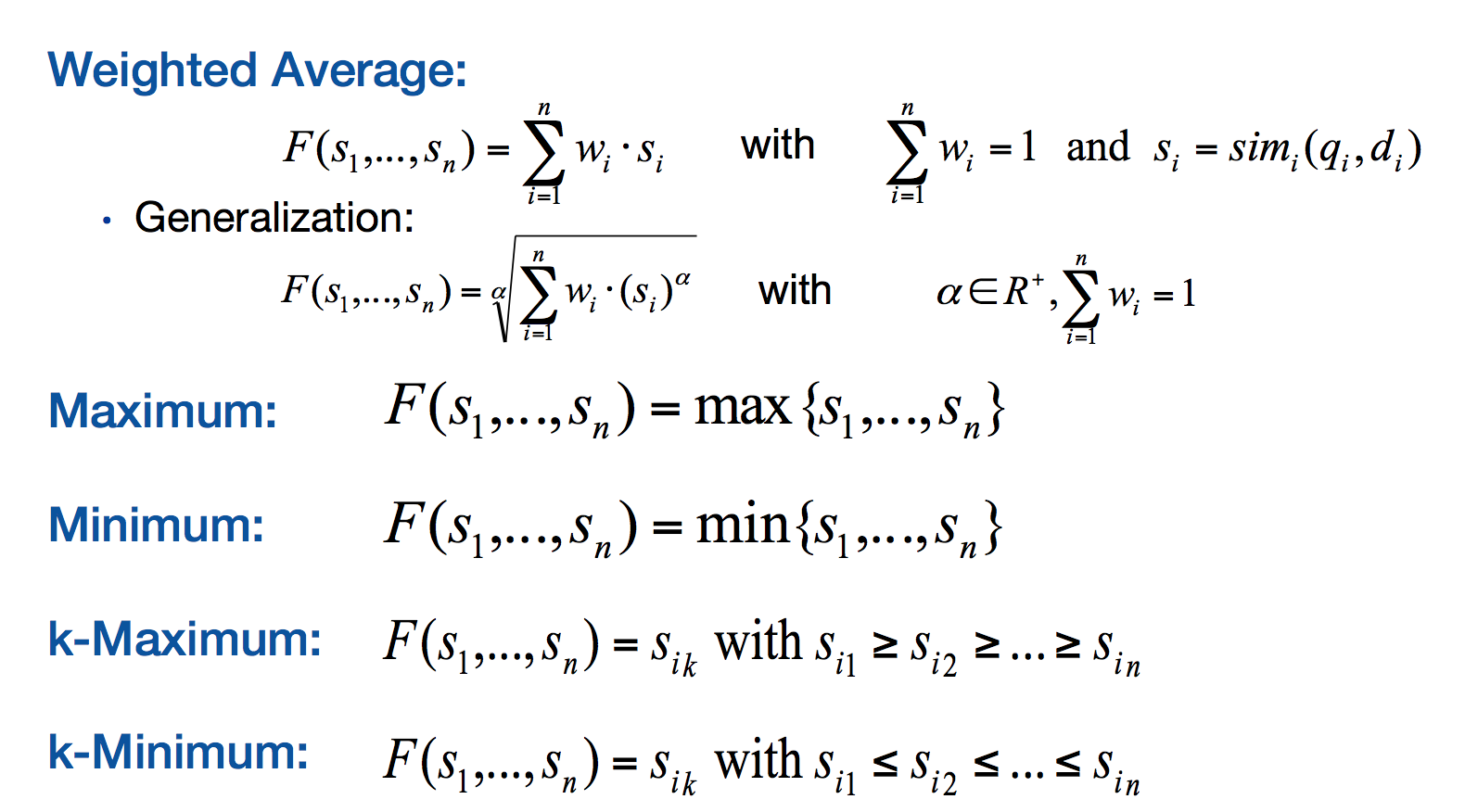
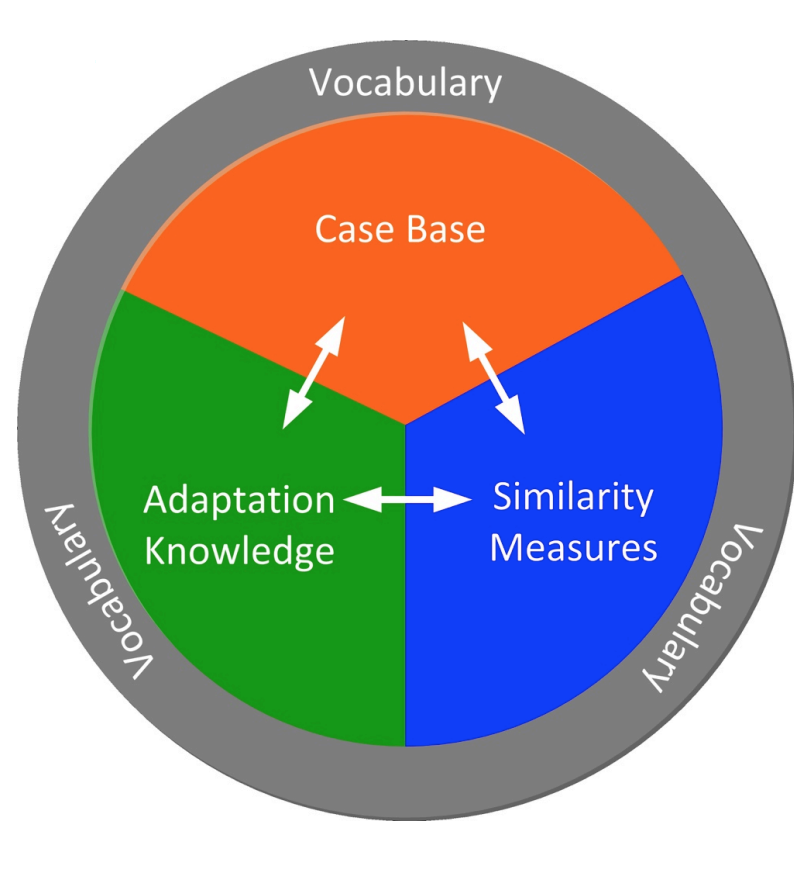
Assignment two

Machine Learning

Sigve Skaugvoll, MIT, H2018

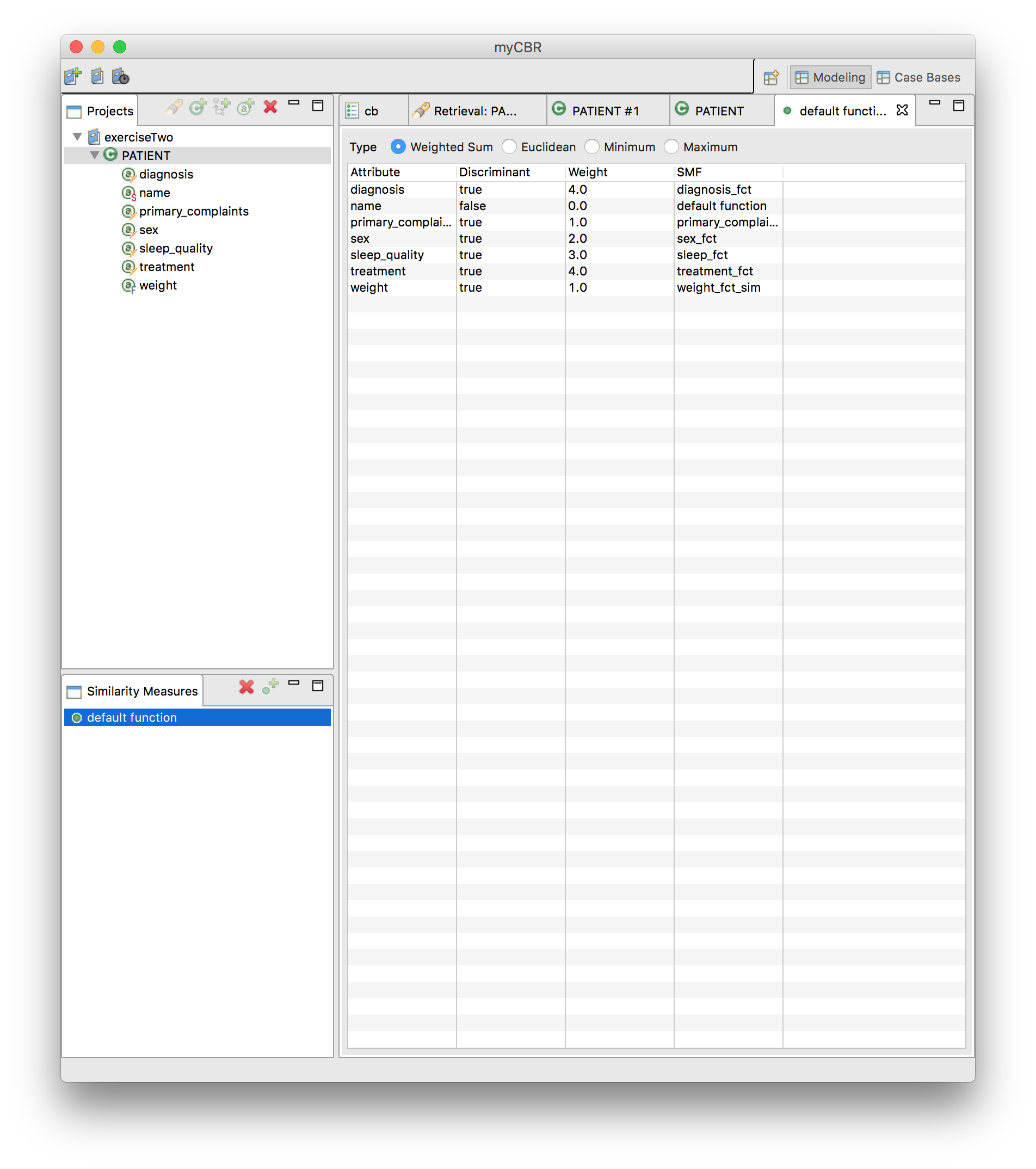
# Theory

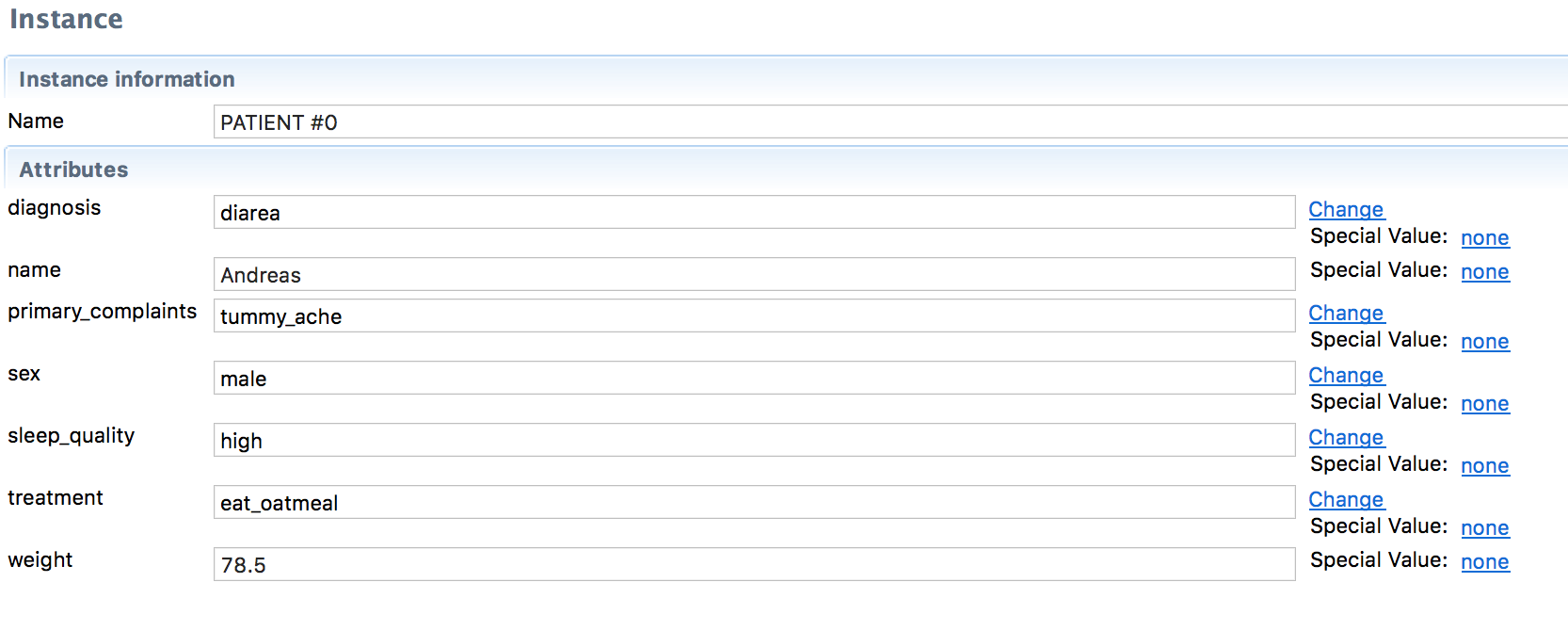
1. **What characterises case-based reasoning (CBR) methods? How are they different from other machine learning approaches?**
   1. Case based reasoning is different because it relies on previous “cases” to find similarity towards the problem at hand. CBR tries to model human problem solving behaviour, because similar problems have similar solutions.  
        
      CBR in the end boils down to structure and methods from CS; Retrieve, Reuse, Revise and Retrain from the general knowledge.  
        
      Another way CBR is different is that it’s a “lazy” learning method. - defer the decision of how to generalize beyond the training data until a new query instance is observed. Instances are typically represented using more rich symbolic descriptions  
        
      Relying on knowledge-based reasoning rather than statistical methods.
2. **Discuss (some of) the ways in which cognitive science has influenced CBR.**
   1. CBR tries to model the way CS has figured out how humans think / reasons to find solutions. Humans tries to find similarity / distance between previous cases and the new problem, the same as CBR. - instance based learning.
   2. CBR uses a form of A script, which is a memory structure that stores general knowledge about typical situation patterns. Such patterns can be MOP and TOP
3. **Methods to evaluate the degree of similarity between two cases are essential in CBR. What is the difference between surface similarity and structural similarity? Give some examples for each approach.**
   1. 
      1. Surface:
         1. Tom chases Jerry and Jerry chases Tom has the same surface-similarity. And the solution to make Tom stop chasing Jerry, would most likely be almost (if not) the same as making Jerry stop chasing Tom.
      2. Structural
         1. Someone gets their car stolen - one solution is to call the police to notify / press charges / get help looking, and someone forgets where they parked, and had their child in the car on a hot summer day. - One solution is to call the police as well to get help since it’s a critical problem. The 2 problems may look different on the surface, but have the same underlying similarities
4. **Explain how the similarity between cases can be measured when cases are made up of attributes with different data types. Give an example of how this can be done.**
   1. We can use the Local-Global principle and apply amalgamation function. This has certain advantages; decomposition of the complex global similarity measure into easy to handle local similarity measures. The drawback is that attribute dependencies have to be considered and modeled with certain “tricks”.   
      
   2. HUSK å gi eksempel
5. **What are knowledge containers in the context of CBR? Give a brief explanation of the different containers.**
   1. Define the facts and the rules in such a way that all the intended knowledge is represented. Then these facts and rules are distributed into the knowledge containers - which then becomes the available knowledge distributed into domains. The containers are dependent of each other do be able to solve a problem.
      1. Consists of these four;
         1. Similarity Measures
            1. The retrieval of similar cases is based upon the use of similarity functions (or measures) to compute the distance or similarity of two cases
         2. Case Base
            1. The systems experience is stored as cases within the case base which can be seen as a special form of a database
         3. Vocabulary
            1. The cases themselves, the similarity measures and the adaptation knowledge are composed upon a vocabulary that contains the objects of interests(terms, attributes, concepts).
         4. Adaptation knowledge
            1. Adaptation knowledge is used whenever a retrieved case’s solution has to be adapted to be suitable to solve the presented problem.

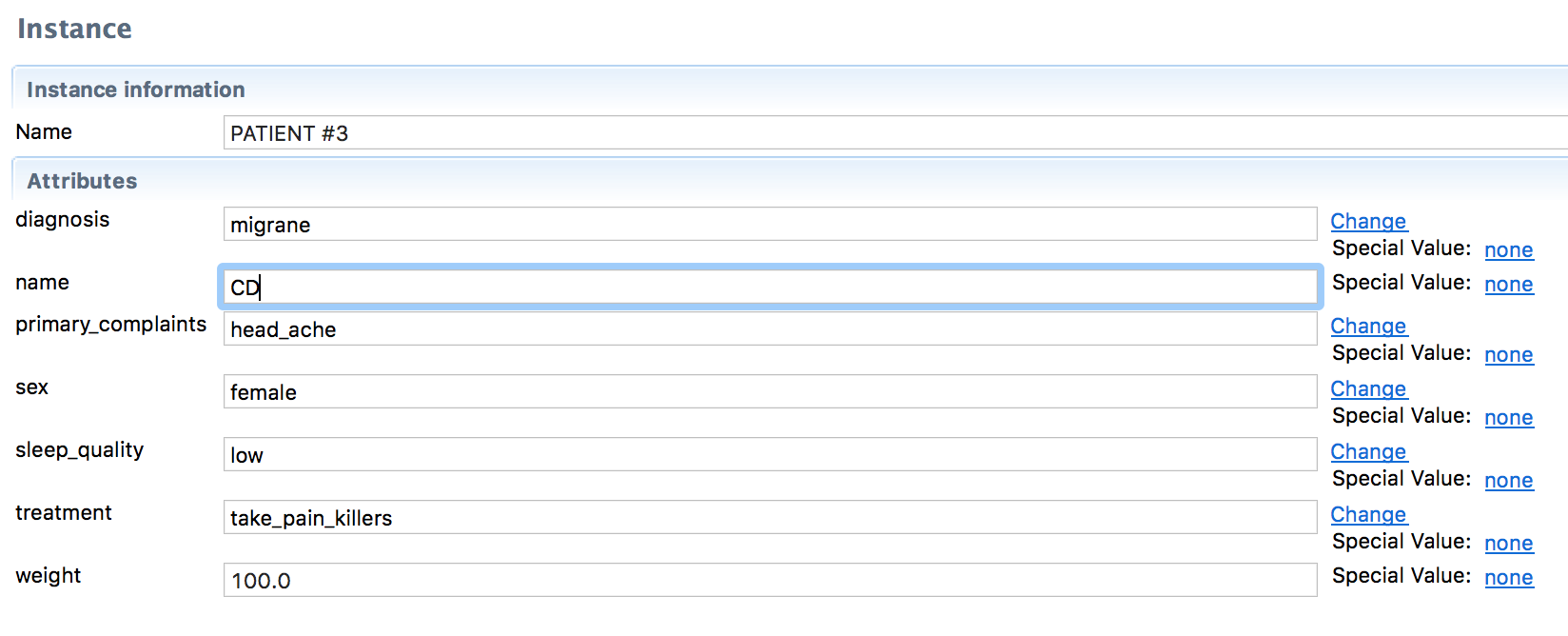
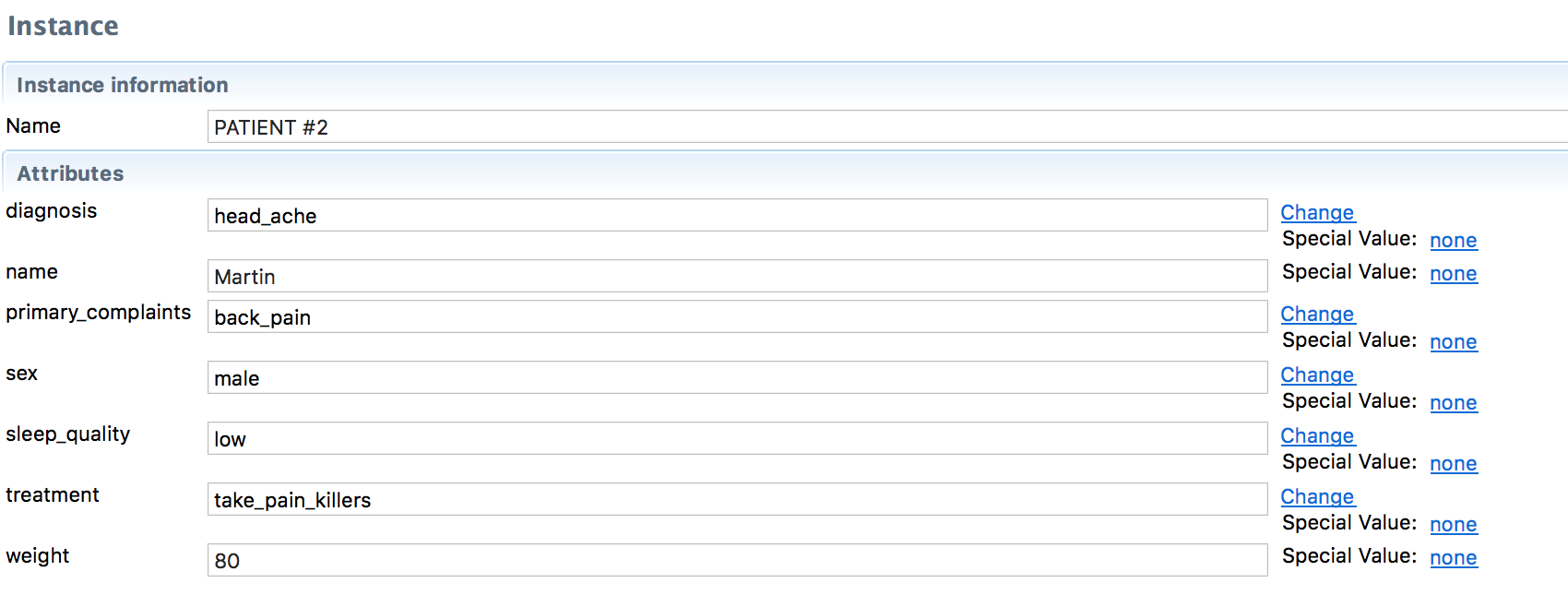


# Practical

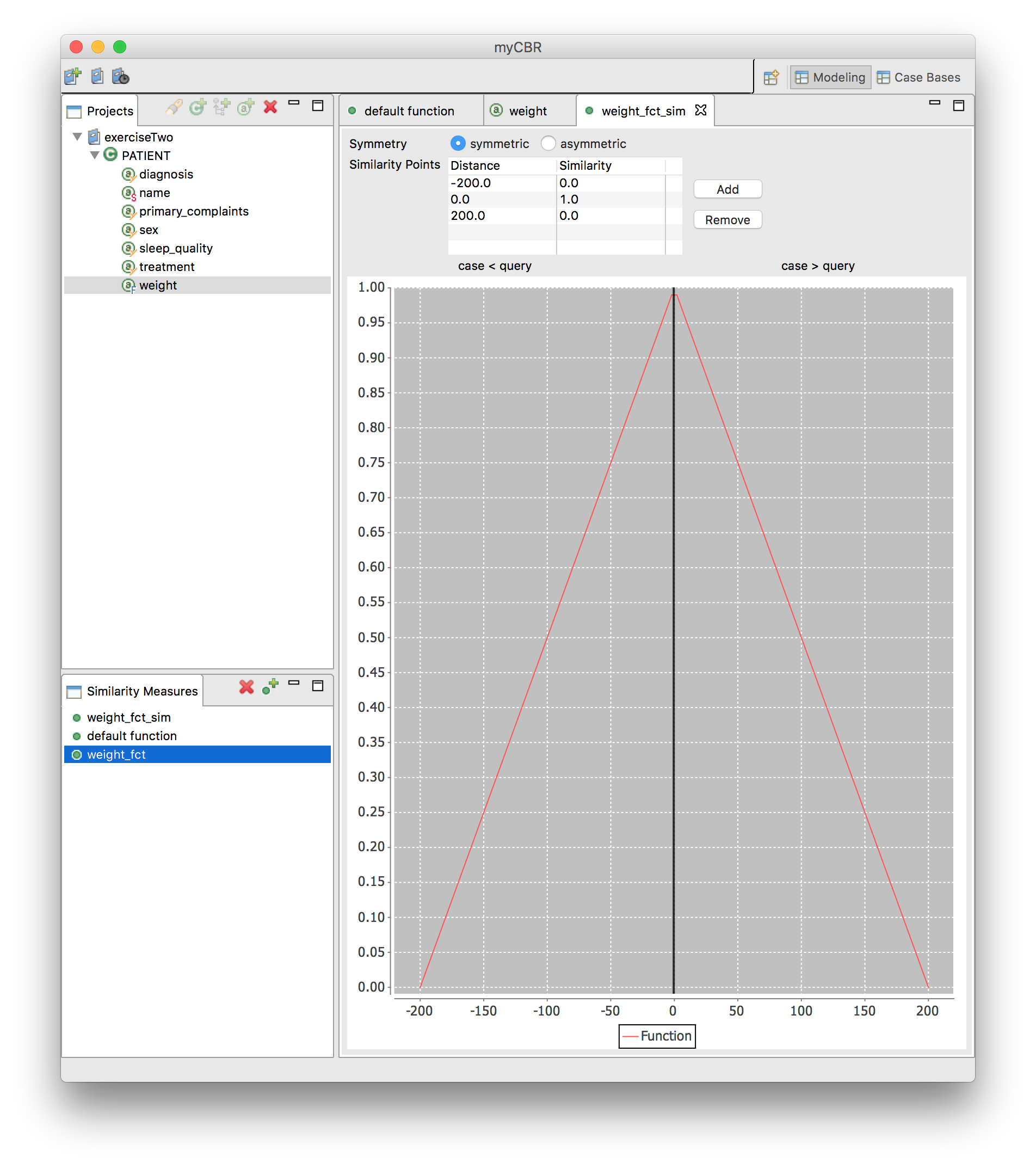
## Case Modelling



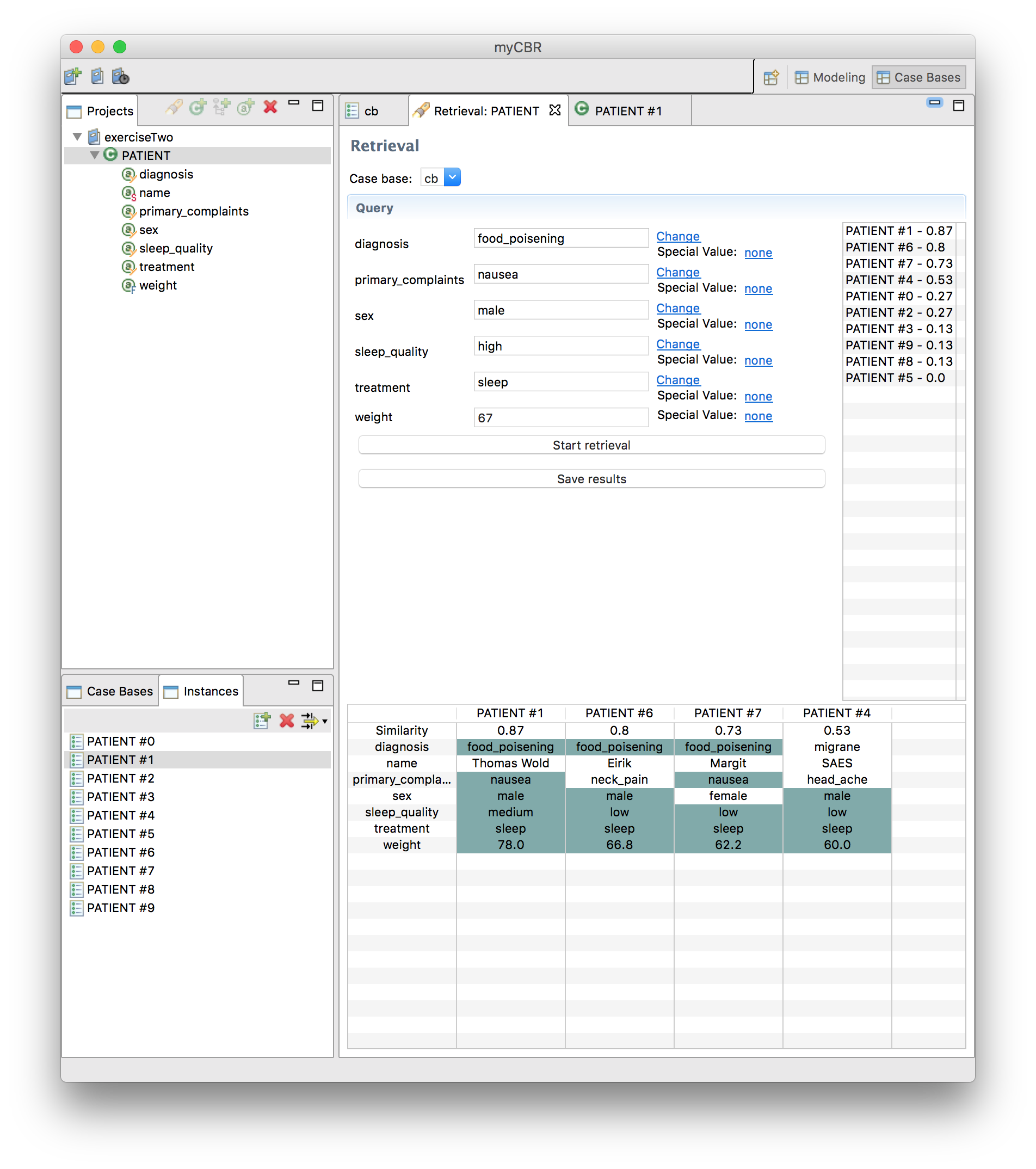


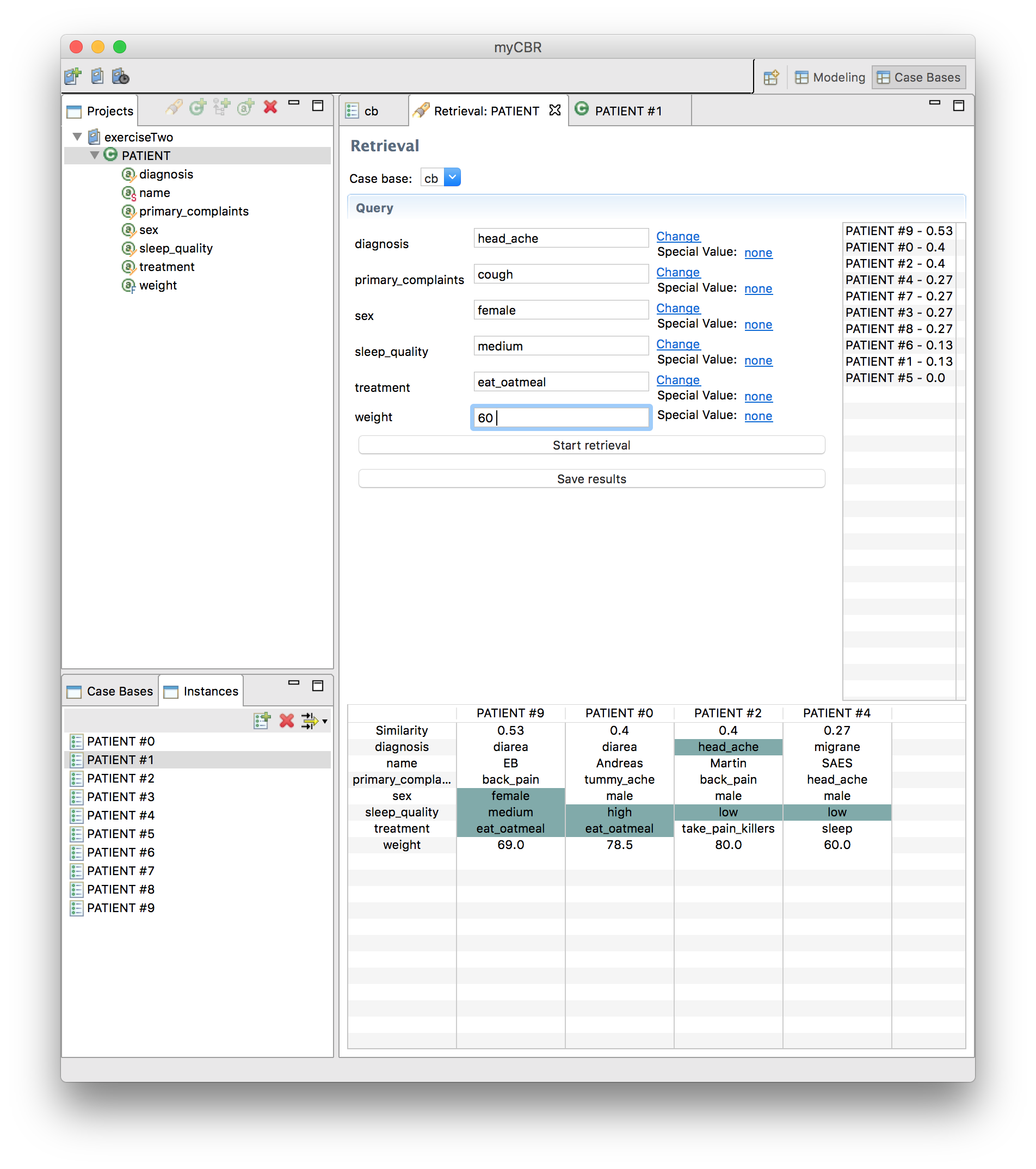


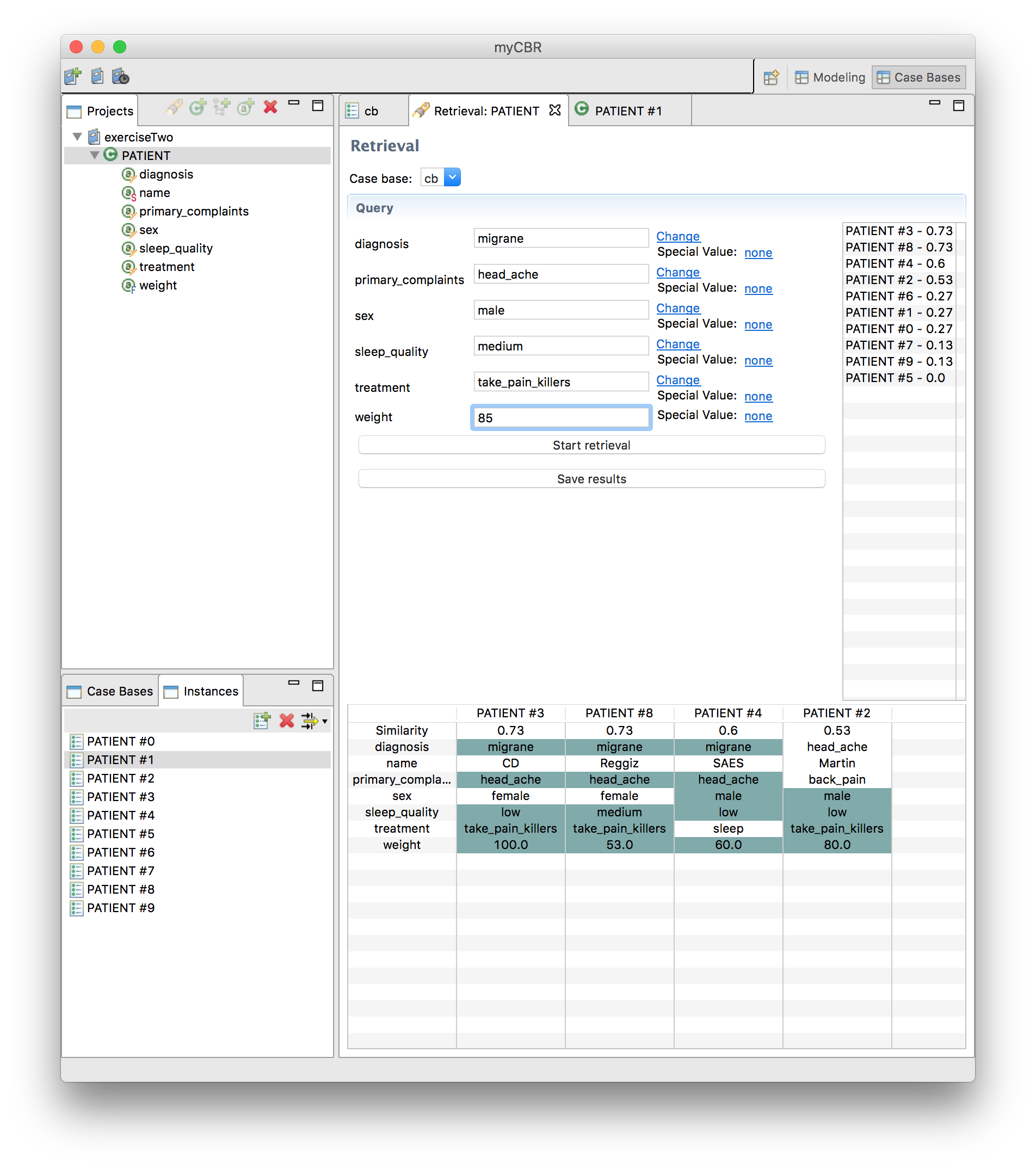
## Case Retrieval

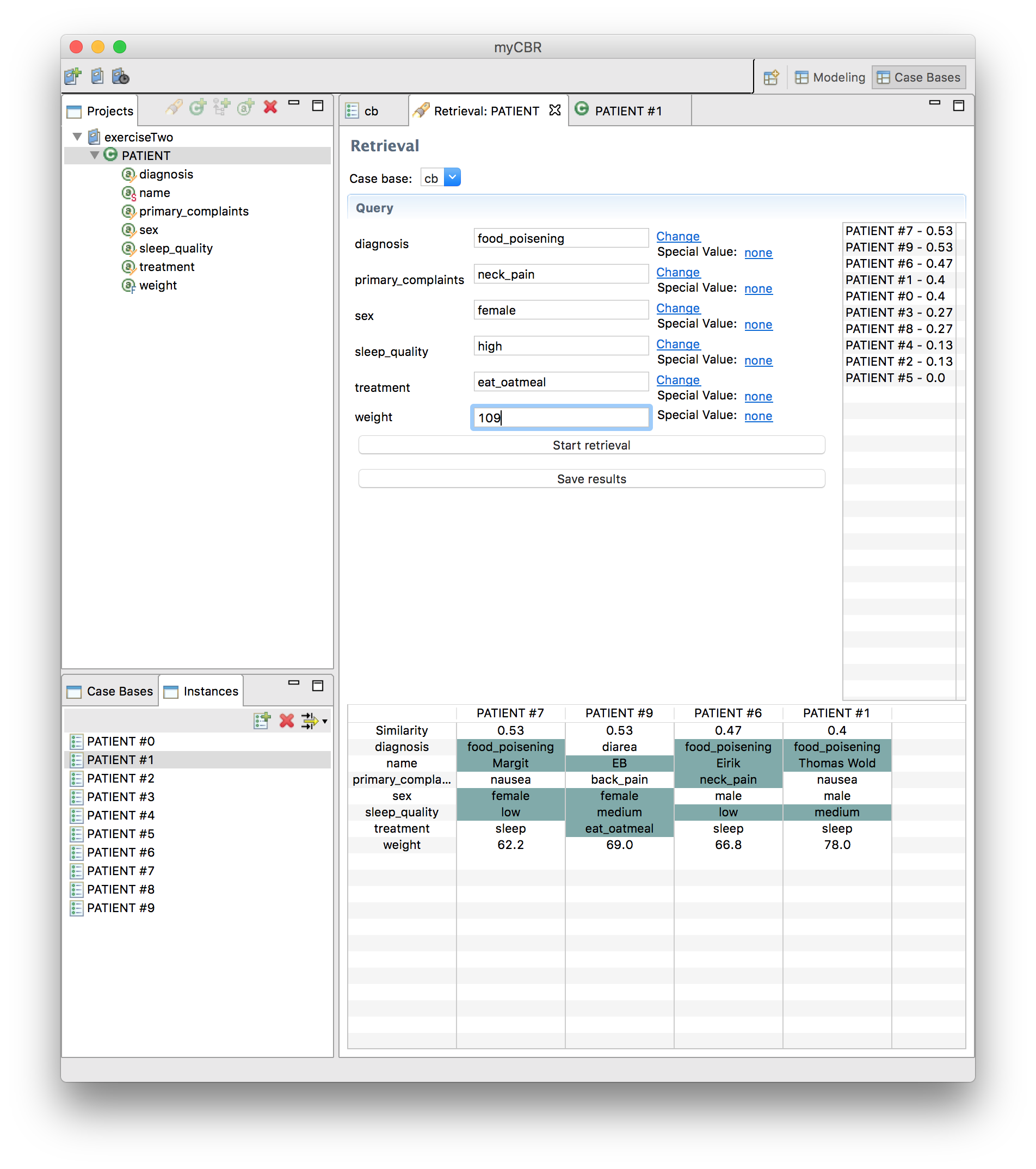


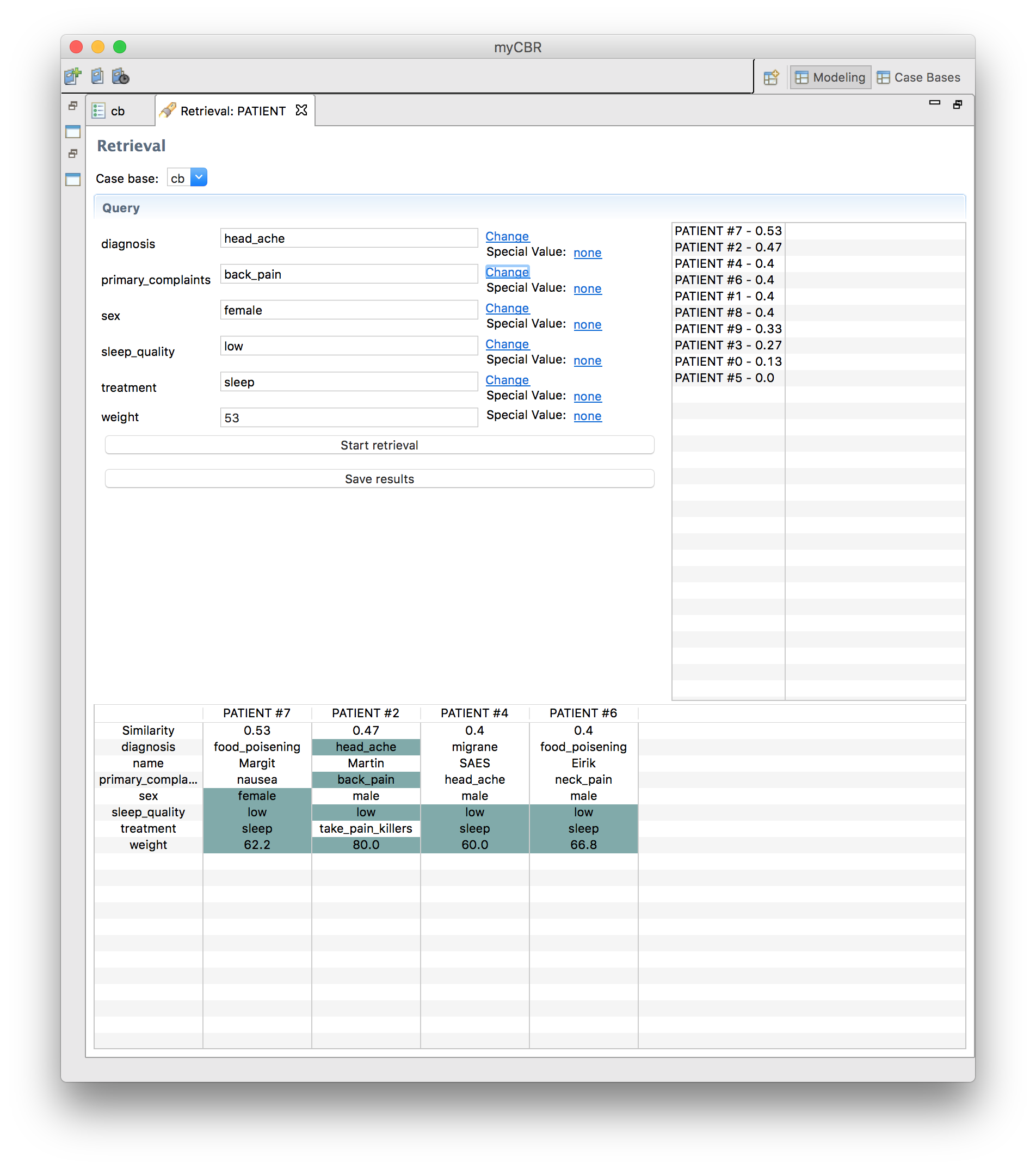
Retrievals:

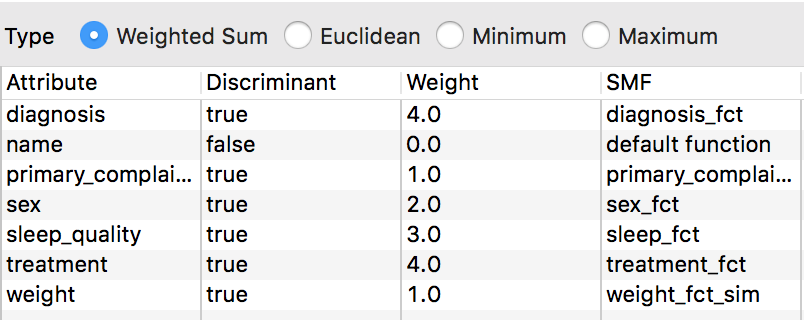






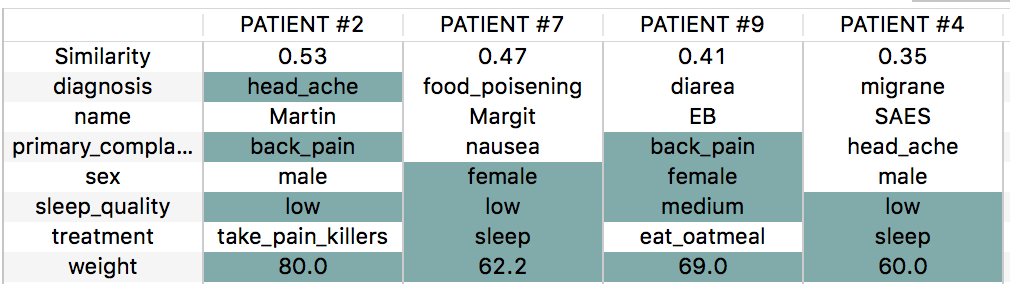






The global similarity has a lot to say! About everything to say, when retrieving. Changing the weights and the “discriminant” changes the retrieval results.

On the last query, the result was interesting! I find it quite interesting that **Food\_poisening** was more similar than **head\_ache**. The reason for this is because the global similarity weights diagnosis and treatment has the same weight. And both patient 7 and 2 has one each of this weight. Sex weights more than primary\_complaint, thus patient 7 has more similarity here.

The weight 62.2 is closer to 53 then 80. Thus patient 7 has more weight, then patient 2, even though “weight” is the same for both. It would make more sense that primary\_complaints has more weight than sex. If we change this set primary\_complaint to 3, and sex stays at 2, and take a look at the new result; We can see that this result gives more meaning because patient 2 and 7 swapped places.

It’s clear that head\_ache should be the most similar case, and weigh primary\_complaints more than sex.

If we compare patient 2 and 9 (first and third place) we see that both patient has back\_pain, patient 9 is closer to the query weight, and is female, but since the diagnosis has more weight, we get the wanted result.

**Q:** The CBR cycle include retrieve, reuse, revise, and retain. Come up with one or more problems that use the patient concept. Briefly explain how each step in the CBR cycle can be executed with the help of myCBR

myCBR can only help with retrieve, there is no reuse,revise and retain functionality. The reuse and steps after is up to the myCBR user to execute. I guess the user can create new instances with the new retained information and create a case that will help improve the system for further use. myCBR is quite hendy for retrieval. It’s quick and easy to create concepts, add attributes, create similarity functions and query.

The system and concept for PATIENT, can be used by an online Artificial intelligence doctor-agent, help for doctors and or supplement for doctors in third world countries where there is an shortage for doctors. Thus the system can “reason” and then subscribe common (non-lethal) drugs for patients.