Subject: Re: non-toxic Build Evolutionary Library
Posted by thomas on Fri, 08 Sep 2023 10:37:21 GMT
View Forum Message <> Reply to Message

I just realize that while fuzzy score and fitness calculation are very similar and partially use the same default values, the function for calculating the score are different. Also in case of fitness it is not obvious how the score is calculates from the real values. Probably, I should use the fuzzy score UI also for the fitness panels. But for the time being, the calculation for the property fitness scores is as follows:

- If the property value is between the desired min and max, then the fitness score is 1.0. If the property is above max or below min, then the score is calculated as score=e^(delta/const) where delta is max-property or property-min, respectively, and const is a property dependent constant that defines how quickly the score decreases from 1.0 when the property departs from the allowed window. Currently, for cLogP const=0.5 and for toxicity const=3.0. This causes a cLogP score 0f 1/e if a cLogP value is 0.5 off the window and a tox score of 1/e if the toxicity indication sum is 5 (assuming that 2 is the allowed upper limit. My current gut feeling says that const should be higher for cLogP and lower for tox risk. In any rate it would be better, if this could be changed from a default value in the fitness panels by the user.

As a reminder: toxicity indication sum is the sum of toxity indation values over all four toxicity classes where a low risk label contributes a 1 and a high risk label contributes a 2. Thus, the maximum would be 8 for a compound that is classified 'high risk' in all four categories.

Once we have the fitness values of all individual properties (0.0 to 1.0), the overall fitness is calculated as a geometric mean of all contributing values considering the defined contribution weight. This is done by calculating the product of all weight*fitness values and then drawing the n-th root from it, where n is the sum of all weight values.

This should answer all the questions I hope.