QCheck Algorithm

Search Algorithm:

Algorithm 1: Correlation

This algorithm depends on computing a metric, a correlation coefficient, that relates an unknown spectrum to each entry in the library of spectra. The entries are ranked in descending order based on the magnitude of the metric. Finally the entries that have the largest metric are reported as most similar to the unknown compound.

The QCheck algorithm differs from the Search algorithm in that the unknown spectrum is modified to match each library or spectrum region and/or resolution but in Searching the unknown spectrum is modified only once to match the set of library spectra.

The algorithm for computing the correlation value between an unknown spectrum and a specific library spectrum or spectrum in a directory is described below.

- Let X be the vector that represents the spectrum of the specific library compound. Then X has the form {x1, x2, x3, …, xn-1, xn} where xi is the intensity at the location indicated by i.

- Let Y be the vector that represents the spectrum of the unknown compound. The length of vector Y is the same as the length of vector X and that X and Y line up.

- Before performing the correlation calculation there is an important step. An estimate of the derivative of X is computed and let's call the result X'. Also the estimated derivative of Y is computed and it is called Y'.

- The derivative that is computed is a three point smoothed derivative. As a result the X vector entries become the X' vector entries according to the rules shown below.

$$x'_1 = 0$$

$$x'_2 = (x_3 - x_1) / 2$$

$$x'_3 = (x_4 - x_2) / 2$$

$$\ldots$$

$$x'_{n-1} = (x_n - x_{n-2}) / 2$$

$$x'_n = 0$$

- The square of the correlation value is computed according to the equation shown below where X′*Y′ denotes the dot product between vectors X′ and Y′. (The dot product gives a scalar.)

$$r^2 = [\ (X'*Y')\ (X'*Y')\ ]\ /\ [\ (X'*X')\ (Y'*Y')\ ]$$

- The metric that is reported is one hundred times the square root of r2. The minimum value that this metric may attain is zero. This value implies that there is no relationship between the unknown spectrum and the particular library spectrum. The maximum value for the metric is one hundred. If the metric is one hundred, then the unknown spectrum is either identical to or the negative of the particular library spectrum.

High Sensitivity:

A Search method includes an algorithm used to generate a metric describing the similarity between the spectrum measured on an unknown material and each spectrum in the library. The displayed results of a search are a list of the materials in the library that are most similar to the unknown material. The metric is also reported for each of the materials in the library. The magnitude of the metric is used subjectively to determine the confidence in a particular match.

Both the QC
of an unkno
known mate
“spectral lib

In the QCheck method the correlation algorithm is always used. The metric that is generated is different, since a monotonic scaling function is applied to the result before it is reported to spread the similar results from each other. The scaling function is described below, where rold is the metric generated in a Search method, and rnew is the metric generated in the QCheck method:

$$x = r_{old} / 100$$

$$r_{new} = 100\ (\ x^{21} + x\ )\ /\ 2$$