3.2.2 Parameterization

A natural way to define sub-instances for the optimal alignment of strings S and T is to optimally align certain substrings of S and T. It turns out that optimally aligning prefixes of S with prefixes of T is sufficient. Thus, for every index i between 0 and n and every index j between 0 and m consider prefixes $S[1 \dots i]$ (empty string for $i = 0$) and $T[1 \dots j]$ (empty string for $j = 0$) and let $\alpha_{i,j}$ be the value of an optimal alignment between strings $S[1 \dots i]$ and $T[1 \dots j]$.

3.2.3 Bellman Principle

If in an optimal alignment of strings $S[1 \dots i]$ and $T[1 \dots j]$ the last column of aligned symbols is deleted, one obtains an optimal alignment of strings $S[1 \dots i-1]$ and $T[1 \dots j-1]$ in case that $S(i)$ was aligned to $T(j)$, an optimal alignment of strings $S[1 \dots i]$ and $T[1 \dots j-1]$ in case that the spacing symbol was aligned to $T(j)$, and an optimal alignment of strings $S[1 \dots i-1]$ and $T[1 \dots j]$ in case that $S(i)$ was aligned to the spacing symbol.

3.2.4 Recursive Solution

Values $\alpha_{0,0}$, $\alpha_{i,0}$, $\alpha_{0,j}$, and $\alpha_{i,j}$, for $1 \leq i \leq n$ and $1 \leq j \leq m$, must be computed. We first treat the case of $\alpha_{i,j}$. Here we observe that at the right end of an optimal alignment of $S[1 \dots i]$ with $T[1 \dots j]$ either $S(i)$ is aligned with $T(j)$, or $S(i)$ is aligned with spacing symbol – , or spacing symbol – is aligned with $T(j)$. We cannot predict which possibility gives the optimal scoring value, thus we must take the maximum of the three values (using here that Bellman's principle may be applied) $\sigma(S(i), T(j)) + \alpha_{i-1,j-1}$, $\sigma(S(i), -) + \alpha_{i-1,j}$, and $\sigma(-, T(j)) + \alpha_{i,j-1}$. For the computation of $\alpha_{i,0}$ we observe that there are no symbols available in $T[1 \dots 0]$, thus there is only a single alignment of $S[1 \dots i]$ with $T[1 \dots 0]$ aligning spacing symbol to every symbol of $S[1 \dots i]$ and having score equal to the sum of all values $\sigma(S(k), -)$, for $1 \leq k \leq i$. Correspondingly, $\alpha_{0,j}$ equals the sum of all values $\sigma(-, T(k))$, for $1 \leq k \leq j$. Finally we have to define basic value $\alpha_{0,0}$. Looking at the recursive formula above we observe that any occurrence of $\alpha_{0,0}$ must contribute value 0 to the overall score. This was also to be expected since the computation of $\alpha_{0,0}$ intends to optimally align empty string with empty string. Thus, we may summarize the recursive computation scheme (usually attributed to Needleman and Wunsch [56] under the name of "global pairwise alignment") as follows: