Problem Set # 1: Due Friday, Feb. 6.
Note Rescheduled Date

This is a hint for Problem Set 1, exercise 5.

5. Counting.
Consider two sequences, \( a = a_1...a_n \) and \( b = b_1...b_m \), where \( 1 \leq n \leq m \).
(a) How many alignments are there of length \( m \)?
(b) How many alignments are there of length \( m + 2 \)?
(c) How many alignments in general are there between \( a \) and \( b \)?
[For a hint on the general case, see p. 19 of Durbin. It is important to note here that for purposes of sequence alignment, one does not allow aligning a gap against a gap. We will, however, distinguish between cases such as

\[
\begin{array}{ccc}
... & a_i & ...
\end{array}
\]

\[
\begin{array}{ccc}
... & b_j & ...
\end{array}
\]

and

\[
\begin{array}{ccc}
... & b_j & ...
\end{array}
\]

\[
\begin{array}{ccc}
... & a_i & ...
\end{array}
\]

Since I don’t require the text by DE, here are the problems on p. 19 of DE:

2.5: Show that the number of ways of intercalling two sequences of length \( m \) and \( n \) respectively, to give a single sequence of length \( m + n \), while preserving the order of each of the two sequences is \( \binom{m+n}{n} \).

2.6 By taking alternating symbols from the the upper and lower sequences in an alignment, then discarding the gaps, show that there is a one-to-one correspondence between gapped alignments of two sequences (with gap-to-gap alignment disallowed as above) and intercalated sequences of the type described in problem 2.5 above.

2.7 Use Stirling’s formula \( (n! \approx \sqrt{2\pi n} n^{n+\frac{1}{2}} e^{-n}) \) to prove that

\[
\binom{2n}{n} \approx \frac{2^{2n}}{\sqrt{2\pi n}}.
\]