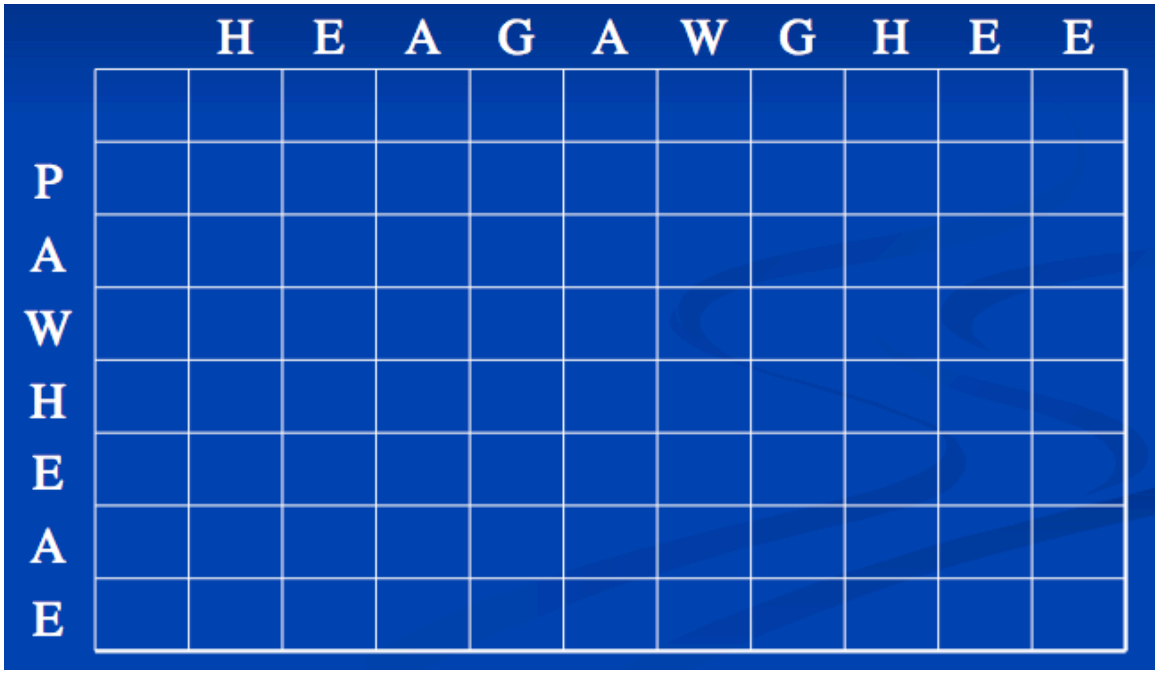
STA/BST226 - Statistical Methods for Bioinformatics
Winter 2007

Exercise: Use the BLOSSUM50 score matrix (Figure 2.2) with a linear gap penalty d=8 to obtain the Needleman-Wunsch optimal global alignment of the sequences HEAGAWGHEE and PAWHEAE. The solution is given in Figure 2.5. Your goal is to fill in the following matrix starting with the initialization condition F(0,0)=0 (top left hand box, above P and to the left of H). Proceed through the top row and left most column using the gap penalty (e.g. F(i,0)=F(0,i)=-id). Finally complete the squares using Equations 2.8, keeping track of which of the three paths was chosen each time you take the max. The optimal alignment is then given by the "traceback" from the lower right hand box (E/E).