Problem

(The solution is found on the next page.) Chiknskratch is a language that uses symbols rather than words. Elizabeth programs a machine that translates an English sentence to a Chiknskratch sentence one word at a time. However, there are several possible Chiknskratch symbols for each English word! Notice below that the word “This” can be translated using 4 different symbols. Elizabeth noticed that different symbols occur next to each other at different rates. For example, “smart chicken” is more common than “intelligent chicken”. She gives scores to word pairs: the higher the score, the more common the word pair is. An English sentence with five words must be translated into five Chiknskratch symbols. In the picture below, arrows labelled with scores connect all valid word pairs for the sentence “This is a simple sentence”. The total score for a translation is the sum of the scores of the four arrows used. What is the highest possible total score for a translation of this sentence?
Solution

The best possible score is 22. This can be achieved by choosing the second, second, second, third, and then finally the first symbol (counting from the top and moving from left-to-right).

How can we find this?
We systematically go word by word from left to right assigning a value to each symbol that is the best possible score for a sentence using that symbol.

To begin, we assign a value of 0 to the four possible translations for “This”. Then we consider the four symbols for “is”. We set the value of each corresponding translation of “This is” to be the maximum label of an arrow coming into the corresponding symbol of “is”. Reading from top to bottom, this gives us values of 3, 6, 7, 4 for the values of translations of “This is”.

To assign values to the translations for “This is a”, we consider the three symbols for “a”. For each symbol, we sum the value on an arrow into the symbol with the corresponding value for “This is” coming from that arrow. Reading from top to bottom, this gives us values of 10, 11 and 13 for translations of “This is a”.

To assign values to the translations for “This is a simple”, we consider the five symbols for “simple”. For each symbol, we sum the value on an arrow into the symbol with the corresponding value for “This is a” coming from that arrow. We then take the maximum of these sums. Reading from top to bottom, this gives us values of 19, 18, 17, 15, and 14 for translations of “This is a simple”.

Finally, to assign values to the translations for “This is a simple sentence”, we consider the two symbols for “sentence”. For each, we sum the value on an arrow into the symbol with the corresponding value for “This is a simple” coming from that arrow. We then take the maximum of these sums. Reading from top to bottom, this gives us values of 22 and 20 for the translations of “This is a simple sentence”.

These “passes” are shown at the end of this solution.

We end with a best possible score for the full sentence is 22. The corresponding translation is “∨ ⊩ ⊂ ≈ ∥”. Using this method to find the best score, we only use 7 + 8 + 7 = 22 additions and some comparisons instead of (20 + 19) × 3 = 117 additions if we checked every possible path through all the nodes. Can you see where these numbers come from?

The Beaver Computing Challenge (BCC):
This problem is based on a previous BCC problem. The BCC is designed to get students with little or no previous experience excited about computing. Questions are inspired by topics in computer science and connections to Computer Science are described in the solutions to all past BCC problems. If you enjoyed this problem, you may want to explore the BCC contest further.

Connections to Computer Science:
The algorithmic idea to solve this problem quickly is called dynamic programming. It is based on a general idea of systematically building the solution from small chunks to bigger and bigger pieces. If you remember (or write down) the partial results, these partial results can be used to calculate a solution without having to recompute these partial results. This problem also gives you a glimpse of contemporary machine translation. It may be somewhat surprising, but machine translation does not depend on a deep understanding of grammar rules. Rather, it works with enormous databases of texts in different languages, and simply put, looks for good matches, especially with digrams and trigrams (pairs or triplets of words that occur frequently).
First Pass:

Second Pass:
Third Pass:

Fourth Pass: