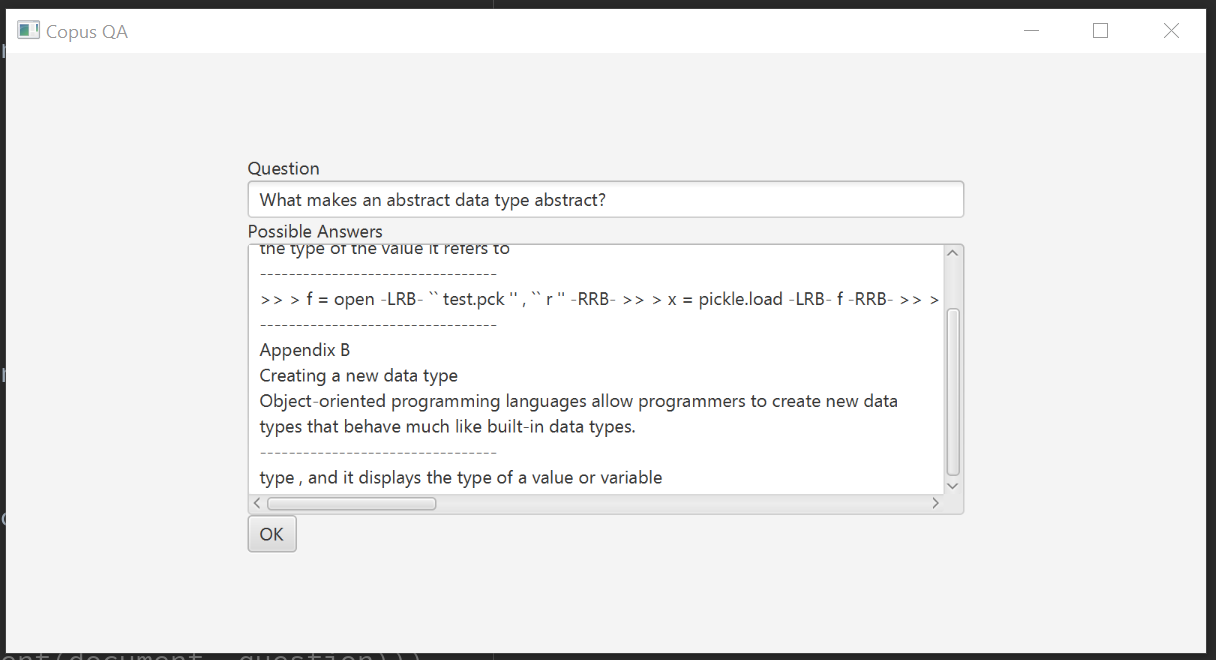
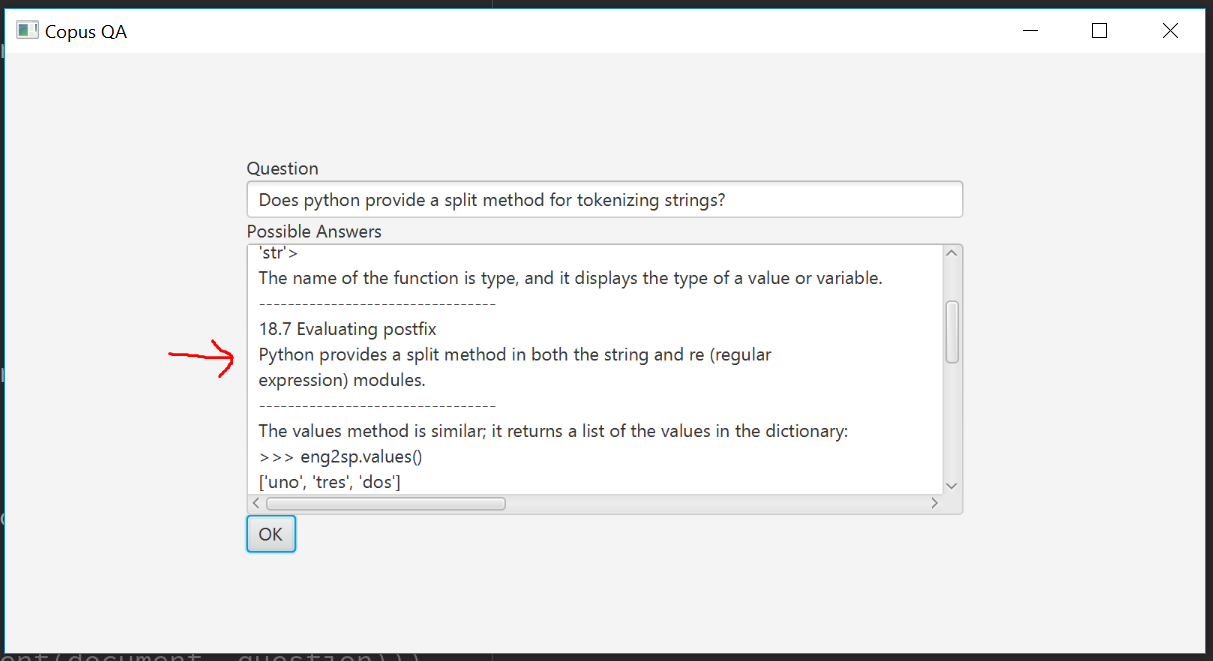


This next one, no answer. The question is too abstract :-)





Next one shows a weakness in that we are processing code as if it were natural language, we need to parse the code in another way, or else recognize and ignore, when we preprocess the corpus

