| Pattern | Function |
|---|---|
| r'\([^)]*\)' | Detect substring within brackets |
| r'[^\(]*\(|\)[^\)]*' | Detect substring outside of brackets |
| r', possibly.*' | Detect everything after given substring |
| r'.*, possibly ' | Detect everything before given substring |
| ' with ' | Used to split string into substrings |
| ' / ' | Used to split string into substrings |

Supplementary Table S2: Patterns to dismantle strings without an exact match into substrings.


| Bing question template |
|---|
| "**Task:** In the following, 6 pairs of terms are listed. Are the terms medically equivalent? Answer only with a "0" if the two terms are not equivalent, a "1" if the terms are medically exactly identical and a "2" if the terms are approximately medically equivalent. Answer with a list of 6 numbers, e.g., "0,1,2,1,0,2", of which the first number is the evaluation of A, the second number is the evaluation of B and so on. \n\n **Pairs of terms:** \nA: "{gpt-4-1st-answer}" and "{correct-answer}" \nB: "{gpt-4-2nd-answer}" and "{correct-answer}" \nC: "{gpt-4-3rd-answer}" and "{correct-answer}", \nD: "{gpt-3.5-turbo-1st-answer}" and "{correct-answer}", \nE: "{gpt-3.5-turbo-2st-answer}" and "{correct-answer}", \nF: "{gpt-3.5-turbo-3rd-answer}" and "{correct-answer}"\n\n **Answer:**" |

Supplementary Table S3: Question template for the evaluation of pairs of medical terms


| Difficulty category | Explanation |
|---|---|
| Very easy | The easiest 20% of all questions. |
| Easy | More difficult than the bottom 20% but easier than the upper 50% of all questions. |
| Intermediate | More difficult than the lower 50% but easier than the upper 20% of all questions. |
| Difficult | More difficult than the lower 80% of questions but easier than the top 5%. |
| Very difficult | The most difficult 5% of all questions. |

Supplementary Table S4: Automatically determined difficulty category by how often it is answered correctly or incorrectly in AMBOSS. All questions are then sorted by difficulty (Source: https://support.amboss.com/hc/en-us/articles/360035679652-Question-difficulty).