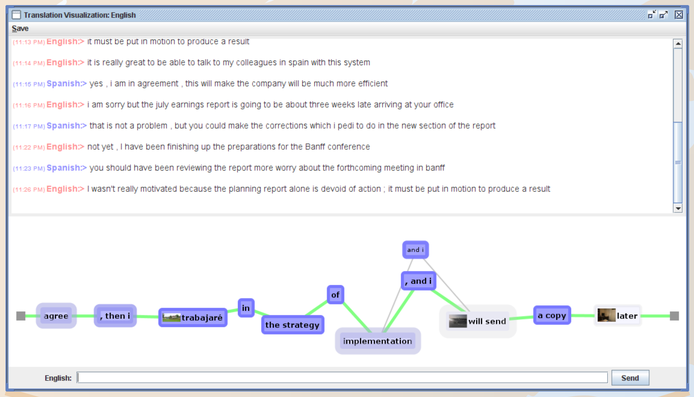
Leveraging Uncertainty Visualization to Enhance Multilingual Chat

Christopher Collins and Gerald Penn

The Potential

Machine translation offers much promise for improving workplace computer-mediated communication among colleagues situated in offices in different parts of the world.

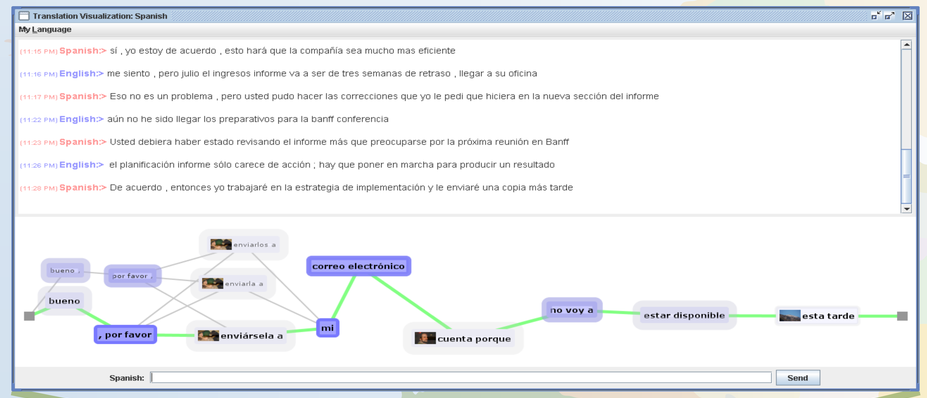


The Problem

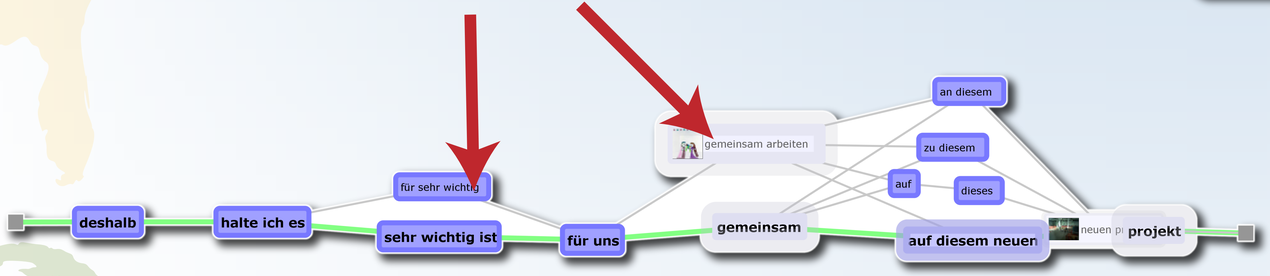
Current translation quality is too low to feasibly use it in mission critical settings. Statistical translation systems generate a ranked list of hypotheses, but all but the first remain hidden.

Our Solution

Our prototype multilingual visualization system for instant messaging conversations reveals the uncertainty in the translation and provides alternative translations when available.

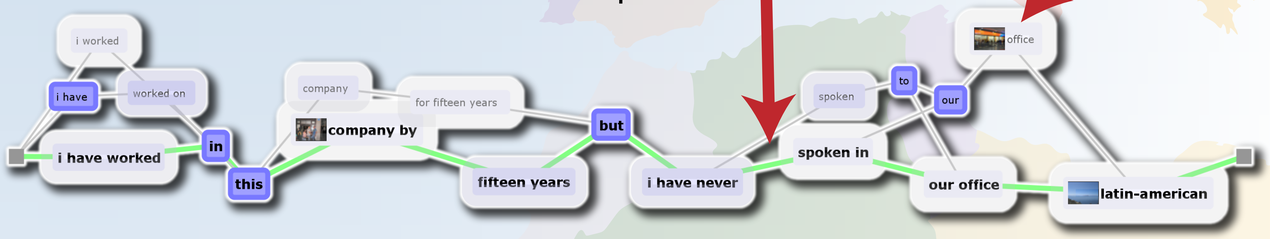


Uncertainty in the translation is visualized through the presence of alternative paths through the lattice. Additionally, the fill colour is directly proportional to the node's confidence score, from saturated blue to desaturated gray. To compensate for variations in colour perception, we redundantly encode the scores in the node border using size and transparency, varying from a tight solid blue, indicating high confidence, to a wide, transparent, gray border, indicating uncertainty.



Green lines reveal the computed best path. Users may click nodes to modify the best path. The final green path is recorded to the transcript.

Photos are retrieved to augment low-scoring translations.



Our bidirectional instant messaging client performs translation on messages it receives using a beam search decoder for statistical phrase-based translation models. The system is trained on 1 million words of the European Parliamentary Corpus for English to Spanish, French, and German.

UNIVERSITY of TORONTO

{ccollins, gpenn}@cs.utoronto.ca

NECTAR

NSERC CRSNG