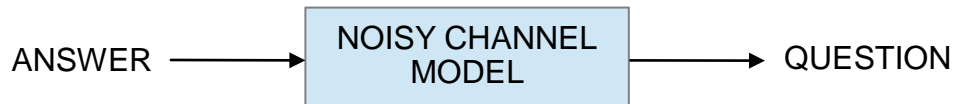


Question Answering Approaches

- Matching same words
 - Does not always work
 - Example: Question: Who is the leader of France?
 - Answer: Hadjenberg, who is the leader of France’s Jewish Community...
- Solution: Some kind of Transformation needed

Noisy Channel Approach

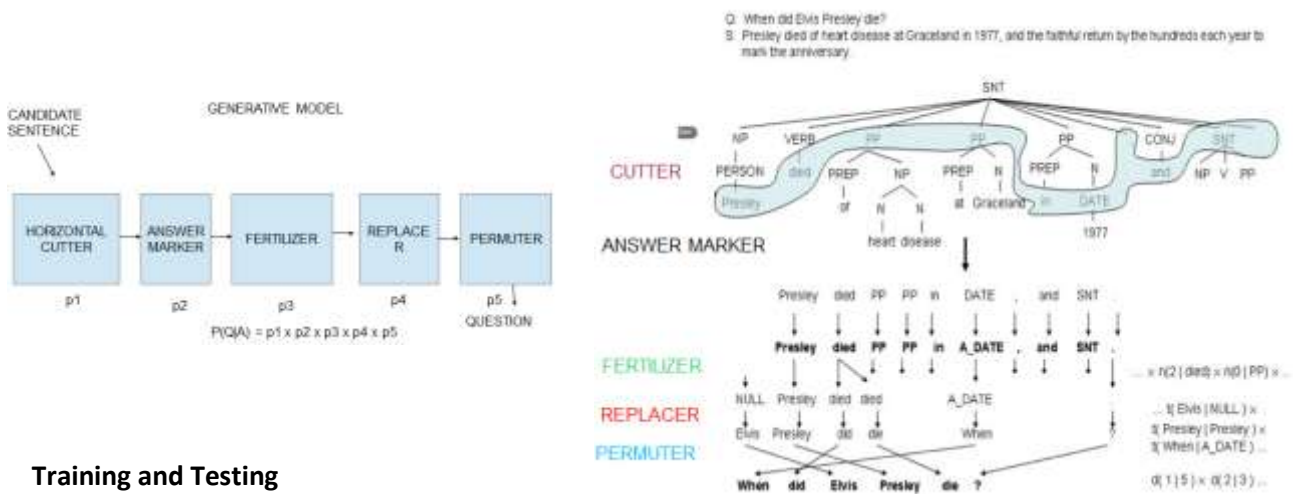
Idea: Question is a noisy version of the Answer



QA System

- IR Engine
 - Get relevant documents and their sentences
- Answer Identification Module
 - From these sentences, identify substrings S(A) as candidate answers
 - For each substring (answer candidate) and sentence, calculate probability of obtaining the Question text using a Generative Model

Answer Identification Module



Training and Testing

- Use the generative model
- Training:
 - Deterministic cutter
 - Answer known
- Testing: Given a Question
 - Exhaustive cutter (try each cut for each sentence)
 - Exhaustive answer marker (try each word as possible answer)
 - Generate probabilities: $P(Q | \text{Sentence} + \text{Cut} + \text{Answer})$
 - Select Sentence+Cut+Answer combination with Max. P

Performance

- Not bad compared to systems with 10s modules, only 2 modules
- Beats QA-base in MRR on TREC datasets
 - (0.325-0.354 vs 0.291)
- QA-base was state-of-art and ranked 2-7 in TREC in last 3 years (in 2003)