LATE2 : MT block: Sample Exam Questions

1. Proposals for machine translation systems make by academics are often accused of lacking the necessary robustness for practical purposes. What does this mean, and what has been done to answer the accusation?
" of the input.

2 Grammatical gender can pose problems for machine translation. Give two or three examples that illustrate this.

3. There is very little interest in interlingual machine translation these days. Why?

4. Two styles of transfer system have been proposed that work by rewriting bags of predicates with order sets of rewriting rules. One was used in the Verbmobil system and the other in an experimental Xerox system. What was the main difference between them?

5. Currently proposed statistical machine translation systems are based on two "models". What are they called, and what do they do.

(sample answers are on the next page)

LATE2 : MT block: Sample Exam Questions With Answers

1.
- A system lacks robustness if, for some inputs, it fails to produce any output at all.
- Current statistical systems are generally robust because they produce the most probably result, however improbable it is.
- Optimality theory has been used to increase robustness, as in Xerox's LFG system.
- Parsers that find no structures for a sentence fall back to producing parses of "fragments" of the input.

2.
- Er ist gross.
  "Er" => "He" or "it"
- "it is on the table"
  "it" => "er", "sie" or "es".
- "What she really wanted to be was a teacher"
  Considerable inference is required to get "Lehrerin" instead of "Lehrer".

3.
- There are many linguistic phenomena that are not well enough understood to allow us to give them a clear meaning such as adjective order, information structure and word order more generally, and compounds.
- It is not clear that there is such a thing as "pure meaning", independent of culture and all linguistic trappings.
- Such a high level of abstraction is probably not needed for practical purposes.

4.
- In the Verbmobil system, predicates were transfer from an input to an output bag as the rewriting process continued so that the output of a rule was never acted upon by a later rule in the sequence.
- In the Xerox system, there was only one bag of predicates so that the output of earlier rules could be input for later rules so that there could be feeding and bleeding among rules as in some phonological systems.

5.
- The translation model finds words (and possibly word sequences) that are likely to occur in the translation of a sentence, given the words (and phrases) that were found in the source.
- The language places these in their most likely order