Overview
History

The field of natural language processing began in the 1940s, after World War II. At this time, people recognized the importance of translation from one language to another and hoped to create a machine that could do this sort of translation automatically. However, the task was obviously not as easy as people first imagined. By 1958, some researchers were identifying significant issues in the development of NLP. One of these researchers was Noam Chomsky, who found it troubling that models of language recognized sentences that were nonsense but grammatically correct as equally irrelevant as sentences that were nonsense and not grammatically correct. Chomsky found it problematic that the sentence "Colorless green ideas sleep furiously" was classified as improbable to the same extent that "Furiously sleep ideas green colorless"; any speaker of English can recognize the former as grammatically correct and the latter as incorrect, and Chomsky felt the same should be expected of machine models.

Around the same time in history, from 1957-1970, researchers split into two divisions concerning NLP: symbolic and stochastic. Symbolic, or rule-based, researchers focused on formal languages and generating syntax; this group consisted of many linguists and computer scientists who considered this branch the beginning of artificial intelligence research. Stochastic researchers were more interested in statistical and probabilistic methods of NLP, working on problems of optical character recognition and pattern recognition between texts.

After 1970, researchers split even further, embracing new areas of NLP as more technology and knowledge became available. One new area was logic-based paradigms, languages that focused on encoding rules and language in mathematical logics. This area of NLP research later contributed to the development of the programming language Prolog. Natural language understanding was another area of NLP that was particularly influenced by SHRDLU, Professor Terry Winograd's doctoral thesis. This program placed a computer in a world of blocks, enabling it to manipulate and answer questions about the blocks according to natural language instructions from the user. The amazing part of this system was its capability to learn and understand with amazing accuracy, something only currently possible in extremely limited domains (e.g., the block world). The following text was generated in a demonstration of SHDRLU:

```
Computer: OK. (does it)
Person: PUT THE LITTLEST PYRAMID ON TOP OF IT.
Computer: OK. (does it)
Person: DOES THE SHORTEST THING THE TALLEST PYRAMID'S SUPPORT SUPPORTS SUPPORT
ANYTHING GREEN?
Computer: YES, THE GREEN PYRAMID.
Person: WHAT COLOR IS IT?
Computer: BY "IT", I ASSUME YOU MEAN THE SHORTEST THING THE TALLEST PYRAMID'S SUPPORT
SUPPORTS.
Computer: RED
```

The computer is clearly able to resolve relationships between objects and understand certain ambiguities. A fourth area of NLP that came into existence after 1970 is discourse modeling. This area examines interchanges between people and computers, working out such ideas as the need to change "you" in a speaker's question to "me" in the computer's answer.

From 1983 to 1993, researchers became more united in focusing on empiricism and probabilistic models. Researchers were able to test certain arguments by Chomsky and others from the 1950s and 60s, discovering that many arguments that were convincing in text were not empirically accurate. Thus, by 1993, probabilistic and statistical methods of handling natural language processing were the most common types of models. In the last decade, NLP has also become more focused on information extraction and generation due to the vast amounts of information scattered across the Internet. Additionally, personal computers are now everywhere, and thus consumer level applications of NLP are much more common and an impetus for further research.