The MITRE Methodology and Survey

Pamela W. Jordan and John W. Benoit
The MITRE Corporation
Bonnie J. Dorr
University of Maryland

During 1991 The MITRE Corp. surveyed and evaluated over 20 machine translation systems across the United States and, to a lesser extent, in Europe and Japan. The intent of the study was to recommend software purchases and R&D support that would address the near-term, medium-term, and long-term translation needs of the three types of users we represented. In this talk we will describe the three types of users and our approach for evaluating a large number of systems in a short period of time (less than one year). When one has to work within these constraints, an approach that is short of complete glass-box and black-box evaluation must be considered. We therefore focused on evaluating systems against user-environment criteria and translation quality needs. A subset of both the glass-box and black-box criteria possible were considered.

To collect the evaluation information we needed, we interviewed developers, researchers, and current users of MT; participated in MT demonstrations; surveyed the literature for additional details about the software; and collected (for further evaluation) sample inputs and outputs for each language handled by the software. Detailed questionnaires were used to guide us in the information collection process. We describe the blind evaluation performed on the collected translations and discuss how we used the collected information to evaluate the systems.