Li Deng
Area Editor, Feature Articles
deng@microsoft.com
www.ee.Columbia.edu/spm

Expanding the Scope of Signal Processing

This special section in this issue of *IEEE Signal Processing Magazine* is dedicated to the theme of spoken language technology, exploring "signal" processing techniques for both speech and language. The articles in this issue cover "language" and/or "understanding" aspects of "signal" processing, beyond the conventional scope of speech processing. This forms a stark contrast to the special section in the September 2005, speech technology in human-machine communication, where only half of the articles dealt with language processing while the focus was on the more traditional tasks of speech analysis and recognition with the trend towards speech understanding.

This change represents a significant new development in our field of signal processing, where the conventional way of thinking about the signal as low-level, numerical-valued information is transcended to a new perspective concerning the importance of high-level, symbolic-valued informational sources (e.g., language or text) that embed their underlying semantic contents. The IEEE Signal Processing Society's Constitution reads: "… The Field of Interest of the Society shall be the theory and application of filtering, coding, transmitting, estimating, detecting, analyzing, recognizing, synthesizing, recording, and reproducing signals by digital or analog devices or techniques. The term 'signal' includes audio, video, speech, image, communication, geophysical, sonar, radar, medical, musical, and other signals…" (article II) While reviewing this, I would naturally classify "language" into one of the nontraditional "other signals" (albeit its prominence as witnessed by the articles in this special section).

Article II of our constitution also lists a range of signal/information processing tasks or application areas, each associated with a particular type of signal. Language as a signal is intimately related to the task of understanding (i.e., discovering the underlying meaning or semantics embedded in the signal), which has not appeared in the traditional list of tasks. Hence, if we expand the traditional

Digital Object Identifier 10.1109/MSP.2008.920380

signals to include language, then we also would expand the traditional signal processing tasks to include understanding.

When we associate each signal type with the full range of the processing tasks, a matrix can be established in which each entry in the matrix represents a specific subfield in signal processing. For example, the row in the matrix for the task of synthesis applies to all types of signals. And for each entry in this row, we would have computer music (for the audio/music signal), speech synthesis (for the speech signal), and computer graphics (for the image signal). After expanding the traditional signals to include language, we will have another entry called natural language generation, a well-established research and application area in computational linguistics and artificial intelligence. Similarly, after expanding the traditional signal processing tasks to include understanding, then we expand the traditional matrix with one new row. For example, the entries in the matrix corresponding to the speech column would contain not only the traditional elements of speech coding, speech transmission, speech analysis, speech enhancement, speech synthesis, and speech recognition but also a new element of speech understanding (or spoken language understanding), which is the focus of the issue in your hands now.

To extrapolate the above idea further, I can propose new expansion of the signal processing tasks to include that of retrieval/mining (a simpler task than understanding). Then, the new row of this further expanded matrix will contain the new elements of music retrieval, spoken document retrieval, image retrieval, video search, and text search (information retrieval). Other possible expansions would include the new task of security and new bioinformatic/ genomic signals. These are not traditional signal processing and as was elegantly pointed out by Prof. Ray Liu earlier, information processing is our destiny beyond the traditional scope of signal processing (see the editorials written in the January and September 2004 issues of this magazine).

Some of the new column elements in the expanded matrix after including language/text in the scope of signal processing are: document summarization (for the task of coding), text parsing (for analysis), spelling/grammar correction (for enhancement), natural language generation (for synthesis), natural language understanding (for understanding), and Web search (for retrieval).

What are practical benefits of constructing the signal/task matrix just discussed? First, we would see a natural expansion of the scope of signal processing with regard to the treatment of language/text as the signal and treatment of understanding and retrieval as the processing tasks, which I outlined above. This would be viewed as the extension of the matrix size in both the row (processing task) and column (signal). Indeed, language adds a new dimension in the signal set with its essential task of content understanding, an ultimate goal of human intelligence. Second, successful machine learning techniques developed for one particular signal type may be more naturally welcomed and examined by researchers working on other signals. Such critical examination would facilitate the potential unification of signal processing methodologies, enhancing the appreciation of and knowledge about the similarities and differences in processing techniques across different signal types. This would be viewed as cross-column propagation from the perspective of matrix construction. Third, we can also benefit from cross-row propagation. For example, while extension from the recognition task to the understanding task is becoming well established for the speech signal (as clearly demonstrated in this special section), are there and should there be similar trends on task expansion for other types of signals?

These are some of my thoughts about this special section. Enjoy reading it. SP