What is “real-world” computing?

Real-world computing will extract meaning from highly complex data. It is important to define precisely what we mean by “complex data”. We distinguish three degrees of complexity in digital data, according to how difficult it is to extract meaning from the data by computer:

1. *Structured text*. This data consists of symbols from a well-defined alphabet organized in well-defined structures, and whose semantics is easily determined from the data itself. Typical examples are spreadsheets and databases. The complete meaning of the data can be internalized by a computer with straightforward parsing techniques.

2. *Free-form text*. This data consists of symbols from a well-defined symbolic alphabet organized in well-defined structures, but whose semantics is not easily determined from the data itself. The semantics may be too complex or may not be well-defined. Determining it may be possible with a lot of computation, or it may be unknown how to determine it. Data mining, which is a branch of machine learning, is a form of meaning extraction from free-form text.

3. *Complex data*. This data is similar to category (2), except that the symbols are more fine-grained and it requires much more processing to determine their semantics. For example, free-form text can consist of Web pages, whereas complex data may consist of digitized photographs or videos. Determining the meaning of part of a photograph, such as face recognition, is much harder than extracting meaning from a textual Web page.

These three degrees of digital data differ mainly in quantitative properties. From degree (1) to (3), they increase successively in abundance and in difficulty of meaning extraction. Existing software mostly handles degree (1); degrees (2) and (3) are mostly transmitted and copied, with creation and meaning extraction done by human users. Some software exists to extract meaning from degree (2): for example, data mining and language translation. Google's corpus-based statistical language translation is an example of meaning extraction from degree (2).

The goal of real-world computing is to address degree (3), which is the most complex but has great potential. Very little software exists for extracting meaning from degree (3), and what software does exist is limited and resource-hungry. Despite this, the processing of degree (3) would be extremely useful for human users, and it will therefore be the focus of this Flagship. Typical examples are real-time audio language translation and assistance for complex real-time human tasks. We assert that learning algorithms are on the verge of being sufficient to successfully tackle degree (3) applications. We motivate this assertion by giving some examples of the state of the art. The popular Dragon NaturallySpeaking speech recognition software handles continuous dictation on a single machine (no pauses between words at fast speaking rates). It is a practical tool to replace typing (at increased speed, since people usually speak faster than they type) [NUA2010]. A second example is Google, which announced in early 2010 that it was working on practical real-time audio language translation [GOU2010]. Prototypes of this have been done much earlier (notably by the German Verbmobil project [WAH2000]). A final example is the IRCAM Research Institute in Computational Acoustics and Music in Paris, France [IRC2010]. IRCAM has developed many industrial-quality tools for sound processing that are important building blocks in a real-time audio translation application.

References

[GOU2010] Chris Gourlay. “Google leaps language barrier with translator phone,” The Times, Feb. 7, 2010 (available on the Web).
[IRC2010] IRCAM (Institut de Recherche et Coordination Acoustique/Musique), www.ircam.fr.
[NUA2010] Nuance Communications. Dragon NaturallySpeaking speech recognition software, www.nuance.com, 2010.
[WAH2000] Wolfgang Wahlster (ed.). *Verbmobil: Foundations of Speech-to-Speech Translation*. Springer-Verlag, 2000.