Preface

Spoken dialog systems have been the object of intensive research interest over the past two decades, and hundreds of scientific articles as well as a handful of text books such as [25, 52, 74, 79, 80, 83] have seen the light of day. What most of these publications lack, however, is a link to the "real world", i.e., to conditions, issues, and environmental characteristics of deployed systems that process millions of calls every week resulting in millions of dollars of cost savings. Instead of learning about:

- Voice user interface design.
- Psychological foundations of human-machine interaction.
- The deep academic[1] side of spoken dialog system research.
- Toy examples.
- Simulated users.

the present book investigates:

- Large deployed systems with thousands of activities whose calls often exceed 20 min of duration.
- Technological advances in deployed dialog systems (such as reinforcement learning, massive use of statistical language models and classifiers, self-adaptation, etc.).
- To which extent academic approaches (such as statistical spoken language understanding or dialog management) are applicable to deployed systems – if at all.

---

[1]This book draws a line between core research on spoken dialog systems as performed in academic institutions and in large industrial research labs on the one hand and commercially deployed spoken dialog systems on the other hand. As a convention, the former will be referred to as *academic*, the latter as *deployed* systems.

To Whom It May Concern

There are three main statements touched upon above:

1. Huge commercial significance of deployed spoken dialog systems.
2. Lack of scientific publications on deployed spoken dialog systems.
3. Overwhelming difference between academic and deployed systems.

These arguments, further backed up in Chap. 1, indicate a strong need for a comprehensive overview about the state of the art in deployed spoken dialog systems. Accordingly, major topics covered by the present book are as follows:

- After a brief introduction to the general architecture of a spoken dialog system, Chap. 1 offers some insight into important parameters of deployed systems (such as traffic, costs) before *comparing the worlds of academic and deployed spoken dialog systems* in various dimensions.
- *Architectural paradigms* for all the components of deployed spoken dialog systems are discussed in Chap. 2. This chapter will also deal with the many limitations deployed systems face (with respect to e.g. functionality, openness of input/output language, performance) imposed by hardware requirements, legal constraints, and the performance and robustness of current speech recognition and understanding technology.
- The key to success or failure of deployed spoken dialog systems is their performance. Performance being a diffuse term when it comes to the (continuous) *evaluation of dialog systems*, Chap. 3 will be dedicated to why, what, and when to measure performance of deployed systems.
- After setting the stage for a continuous performance evaluation, the logical consequence is trying to increase system performance on an ongoing basis. This attempt is often realized as a continuous cycle involving multiple *techniques for adapting and optimizing* all the components of deployed spoken dialog systems as discussed in Chap. 4. Adaptation and optimization are essential to deployed applications because of two main reasons:

  1. Every application can only be suboptimal when deployed for the first time due to the absence of live data during the initial design phase. Hence, application tuning is crucial to make sure deployed spoken dialog systems achieve maximum performance.
  2. Caller behavior, call reasons, caller characteristics, and business objectives are subject to change over time. External events that can be of irregular (such as network outages, promotions, political events), seasonal (college football season, winter recess), or slowly progressing nature (slow migration from analog to digital television, expansion of the Smartphone market) may have considerable effects on what type of calls an application must be able to handle.

Due to the book's focus on paradigms, processes, and techniques applied to deployed spoken dialog systems, it will be of primary interest to speech scientists,

voice user interface designers, application engineers, and other technical staff of
the automated call center industry, probably the largest group of professionals in
the speech and language processing industry. Since Chap. 1 as well as several other
parts of the book aim at bridging the gap between academic and deployed spoken
dialog systems, the community of academic researchers in the field is in focus as
well.

New York City   *David Suendermann*
February 2011

Springer