4. Graphical Models in Speech and Language Research

Jeff Bilmes, University of Washington
http://ssli.ee.washington.edu/~bilmes

Graphical models (GMs) are a general statistical abstraction that can be used to describe a wide variety of problem domains. Recently, significant research has occurred on their application to speech and language processing. GMs offer a mathematically formal but widely flexible means for solving many of the problems encountered in these fields. Because of their generality, GMs make it possible to rapidly go from novel idea to working implementation. In this advanced tutorial, we will survey how GMs can be used to represent structures and models in speech and language.

We start with concepts and notation, including an inspection of different forms of graphical models, and some intriguing constructs these forms make available. This includes the notion of a "switching network", where one portion of a network might determine the existence of another, "sparse dependencies", where many combinations of variable values are forced to have zero probability, and "child observations", where influence can flow in the opposite direction of a directed edge in a graph. We will in general see how GMs can be viewed as a mathematically formal visual language, offering a precise set of primitives for specifying statistical systems. We will continue with an analysis of algorithms for performing probabilistic inference on graphs, concentrating on both theory (e.g., when is inference tractable) and practice (data structures and implementation). We will give special attention to the challenges that arise when the underlying domain is temporal.

Next, we will examine the ways GMs can represent speech and language. This will include explicit representations of hierarchical and temporal phenomena such as parameter sharing, multi-stream models with varying degrees of asynchrony, and classifier combination. We will see how these can be used to represent speech evolution in terms of both phonology and articulation. We will also cover graphical representations of language, including explicit structures for N-grams, interpolation, skipping, hierarchical classes, smoothing, back-off, factored representations, and other forms. Furthermore, we will investigate how to describe statistical machine translation via novel multi-dynamic graph representations.

While graphs not only can represent many well-known statistical models, with only minor graph adjustments they can also represent very different (and potentially novel) systems. We will observe how deterministic dependencies, switching networks, and child observations greatly facilitate this phenomenon. Moreover, we will see how a graph's associated inferential machinery can shield a user from needing to "reinvent the wheel" each time it is desired to investigate a new model.

Lastly, we will briefly survey available GM toolkits and their features. We will include a comparison of GM technology with its modern alternatives. Tutorial attendees will thus learn not only how to use GMs, but also how to decide when and where GM technology is best applied.

4.1. Tutorial Outline

1. Overview and Motivation.
2. Different GM types, constructs, and structures.
3. Theory and practice of probabilistic inference in Dynamic GMs.
4. Explicit representations of temporal structures.
5. Graphical models of speech.
6. Graphical models of language.
7. Graphical models of statistical machine translation.
8. GM Toolkits.
9. GM technology vs. its alternatives.

4.2. Target Audience

This tutorial will assume a basic knowledge of standard language and speech processing, including knowledge of hidden Markov models, maximum entropy models, and the many techniques that go into making such models successful. It will also be assumed that the audience is comfortable with basic statistical terminology.

Jeff A. Bilmes is an Assistant Professor in the Department of Electrical Engineering at the University of Washington, Seattle (adjunct in Linguistics and in Computer Science and Engineering). He co-founded the Signal, Speech, and Language Interpretation Laboratory at the University. He received a masters degree from MIT, and a Ph.D. in Computer Science at the University of California in Berkeley. Jeff is an author of the graphical models toolkit (GMTK), and was a leader of the 2001 Johns Hopkins summer workshop team applying graphical models to speech and language. His primary research lies in statistical graphical models, speech, language and time series processing, human-computer interfaces, and probabilistic machine learning.