

# Predicting Dialogue Acts from Prosodic Information

Sergio Coria, Luis Pineda

Institute for Applied Mathematics and Systems (IIMAS)  
Universidad Nacional Autonoma de Mexico (UNAM)  
Circuito Escolar S/N, Ciudad Universitaria, Del. Coyoacan  
04510 Mexico, D.F. Mexico  
coria@turing.iimas.unam.mx, luis@leibniz.iimas.unam.mx

**Abstract.** In this paper, the influence of intonation to recognize dialogue acts from speech is assessed. Assessment is based on an empirical approach: manually tagged data from a spoken-dialogue and video corpus are used in a CART-style machine learning algorithm to produce a predictive model. Our approach involves two general stages: the tagging task, and the development of machine learning experiments. In the first stage, human annotators produce dialogue act taggings using a formal methodology, obtaining a highly enough tagging agreement, measured with Kappa statistics. In the second stage, tagging data are used to generate decision trees. Preliminary results show that intonation information is useful to recognize sentence mood, and sentence mood and utterance duration data contribute to recognize dialogue act. Precision, recall and Kappa values of the predictive model are promising. Our model can contribute to improve automatic speech recognition or dialogue management systems.

## 1 Introduction

A dialogue act tag characterizes the type of intention which a speaker intends to express in an utterance. A listener has to analyze the utterance, its intonation and its context to identify the correct dialogue act which his interlocutor wants to communicate. Two models to analyze dialogue acts are DAMSL (Dialogue Act Markup in Several Layers) [1] and DIME-DAMSL [2]; the latter is a multimodal adaptation of DAMSL to the DIME project [3]. The Verbmobil Project [4] developed another dialogue act model, which has been used in practical dialogue systems.

DAMSL assumes that dialogue acts occur on four dimensions: communicative status, information level, forward and backward looking function. The communicative status determines if an utterance was uninterpretable or abandoned or if it expressed a self-talk. The information level classifies utterances according to whether they refer to the task, the task management, or the communication management. The forward looking function identifies the effect which an utterance has on the future of the dialogue; this includes statements (assert, reassert), influencing an addressee future actions (open option, action directive), information requests, committing a speaker future actions (offer, commit), conventional (opening, closing), explicit performative, or exclamation. Backward looking function indicates the way an utterance relates to one or more previous utterances; this includes agreement (accept, accept part, maybe, reject