

I RICONOSCITORI MULTILIVELLO E LA GRANULARITÀ TEMPORALE: ASPETTI TECNOLOGICI E ASPETTI COGNITIVI

Francesco Cutugno, Gianpaolo Coro

Dipartimento di Scienze Fisiche – gruppo NLP – Università di Napoli “Federico II”
{cutugno,coro}@na.infn.it

RIASSUNTO

In questa comunicazione presenteremo alcune considerazioni sulla progettazione di riconoscitori di parlato spontaneo basati sulla individuazione di unità linguistiche di dimensioni superiori al fono.

Presenteremo una revisione della scarsa letteratura disponibile sia nel settore applicativo-tecnologico che in quello degli studi cognitivi sul riconoscimento del parlato nell'uomo. In questi studi si osserva come, negli ultimi tempi, stia emergendo un approccio di tipo “multilivello” in cui si ipotizza che differenti processi accettino in ingresso contemporaneamente il flusso di informazione linguistica (o acustica) sottoponendolo contemporaneamente a più processi analitici in parallelo. L'obiettivo di questi modelli è quello di tentare di sfruttare quanto più possibile la ridondanza contenuta nel codice linguistico-informativo.

Nel settore degli studi psicolinguistici sono sempre più frequenti le teorie che prevedono l'intervento di più moduli cognitivi paralleli ed integrati che analizzino l'informazione linguistica a più livelli e che suggeriscono ipotesi su come i differenti livelli possano integrarsi fra loro. Nella nostra comunicazione opereremo una panoramica su queste teorie.

Allo stesso tempo, sul piano tecnologico-applicativo, un primo approccio all'integrazione dei livelli linguistici di diversa granularità può essere osservato in alcuni lavori nei quali viene presentata l'integrazione di due riconoscitori, uno di tipo “sillabico” e l'altro “fonetico” e che portano a prestazioni migliori rispetto a quelle ottenute se i due riconoscitori lavorano singolarmente. Altri riconoscitori basati solo su unità sillabiche, senza fare uso di segmenti di segnale a differente granularità, sono recentemente apparsi in letteratura. In questo quadro di riferimento, anche il nostro gruppo ha sviluppato un riconoscitore automatico basato su unità linguistica di lunghezza sillabica definite in base al pattern energetico del segnale verbale. A partire da questi primi tentativi, dunque, incomincia ad affermarsi la necessità di introdurre in qualche modo all'interno del riconoscitore l'informazione proveniente da segmenti più ampi della sillaba, o addirittura da unità soprasegmentali, che quasi sicuramente intervengono anche quando un essere umano opera un riconoscimento. La nostra comunicazione riporterà una rassegna delle problematiche “tecnologiche” finora esposte. In conclusione presenteremo una proposta implementativa che cerca di integrare i diversi livelli linguistici, operando un riconoscimento su più scale di granularità facendo uso di un modello basato sugli Statecharts. Questi sono evoluzioni della caratterizzazione degli automi finiti, che introducono l'utile concetto di gerarchia di stati, sfruttabile per creare una struttura piramidale per i diversi “granuli” del linguaggio. Scorrendo su è già la gerarchia di livelli per risolvere i casi di incertezza, un sistema automatico potrebbe operare un riconoscimento molto più efficace. Il modello proposto in questa sede è un prototipo che mostra un modo per integrare i diversi livelli in maniera da sfruttare le qualità di differenti ASR, che si rivolgono a distinte scale di granularità.