

ABSTRACT

Text to Speech is one of Information Technology application which is used as an interaction between human and computer by converting text becoming voice. Nowadays, TTS Bahasa Indonesia has been made for PC use that is IndoTTS, but IndoTTS reading is still unnatural.

In Text to Speech system, a forming of the right intonation is very important factor that will influence voice in system output. Part of this system that arranges intonation forming in output part called 'prosody evocation'. Intonation, resulted from this part, includes duration and pitch of input text pronunciation.

Process of duration value determination and pitch from input text have complex and non-linear characteristic, so as a base of prosody model system Multilayer Perceptron Neural Network (MLPNN) is used. Prosody model based on MLPNN produces intonation of input text pronunciations by deciding duration value and pitch from every phoneme of input text compiler. Determination of duration value and pitch is done after system does some learning about pronunciation sample from a sentence.

In this final task, Text to Speech of Indonesian language application has been made and accompanied by natural improvement in reading a sentence without numeral and symbols. This application is done by using programming language Borland Delphi 7.0 with diphone Indonesian language database which has been available and MBROLA pronunciation generator. After getting the MOS result from 30 coresponden, it can be concluded that TTS system with MLPNN prosody model has better quality compared with IndoTTS system.

Keyword : Prosody, MLPNN, diphone, MBROLA.