ISE 599 paper review
Title: The dynamics of dynamics: A model of musical expression
Author: Neil P. McAngus Todd
Reviewer: Jie Liu

This paper describes a model of musical dynamics in a very concise manner. First, the definitions of tempo and dynamics/intensity/loudness/volume are given. Tempo could be understood as velocity, and intensity could be understood as a kind of energy. So they could be modeled or mapped to variables in simple physical motion models, which the author did in the latter part of the paper.

Then the author gave the idea that the crescendo/decrescendo pairs could mark a phrase. This idea was discussed by an example, 8 bars in Mozart's A-Major Sonata K331, and supported by some previous research. Since crescendo/decrescendo pairs represent the change of the dynamics and could mark musical phrases, we can infer that the change of dynamics could be linked to musical structure, which is the basis for musical phrasing. Actually this proposition is also mentioned by the author in the latter part. Moreover, from the Fig.1, the author gave another observation, the more important the boundary, the greater the decrescendo/ritardando.

Tempo is coupled with dynamics, which could be observed from some performances. The faster the tempo, the louder the performance. Two graphs (Fig.2) are shown in the paper to support this observation. Till now, the author has shown all of the 3 propositions, which are that: (1) crescendo/decrescendo marks musical phrases, (2) the shape of the phrases is a function of structural importance, and (3) tempo is coupled with dynamics.

Todd claimed that there are 6 components for a model for musical expression: a representation of structure, a performance procedure, an encoding function, an independent variable (metrical distance or time), a set of structure variables and a set of style parameters. I admit that those 6 components are necessary for the AES framework, based on his previous work. But, I argue that if we are not going to use a physical motion model to model tempo or dynamics, those components would not all be necessary. I would have to read more papers on music expression modeling to study this topic in greater depth.

The next part in the paper is the physical motion model used to derive the encoding function for intensity, using the metrical distance as the independent variable. The process is straightforward, but contains some approximations and assumptions. The coupling of the structure variable is a little confusing. It would have helped to have an example.

Then the author talked about the model validation. The performance data is the input, and then he listens to it, guesses the structure variable, and then sends to the model as input. The output of the model, the simulation, is compared with the original performance data, to do some modifications. He runs the whole process recursively until the difference between the simulation and the original data is acceptable.

I doubted that whether it was worthy to model musical expression before, because I think it is impossible to model such an individual and creative process. However, after reading this paper, I realized that it is still good to explore some common rules inside the musical expression, although it appears to be difficult. For this paper, which was written more than 10 years ago, the work shown does have some limitations. For example, the model was not tested with more examples so we can question its generality. But I still appreciate this work, because it presents a method to get the facts related to musical expression, such as musical structure, tempo, intensity, phrasing and so on. It provides a base on which we can change or improve, to possibly model expression.