Continuation of a Path Diagram to Syntax Application

Background & Motivation

Structural equation modeling (SEM) is a statistical technique that is becoming increasingly popular in the educational and behavioral sciences. SEM allows researchers to test the validity of hypothesized models involving complex relationships among multiple variables. Collected data is used to estimate the parameters of the equations and assessing the fit of the model.

There are several popular commercially applications for conducting SEM analyses, but all have serious shortcomings. Two of the most powerful applications, MPlus and EQS, require models be detailed in syntax rather than visually. AMOS™ does allow the user to create path diagrams by dragging and dropping the desired elements, however the interface is antiquated and it is often difficult to achieve the desired results. In addition, all three software packages are available only for Microsoft Windows® operating systems and academic licenses cost several hundred dollars apiece.

An alternative to the commercial statistical software packages is R, a free software environment for statistical computing. R runs on UNIX platforms, Windows and MacOS (R Foundation, 2010), two structural equation modeling packages *sem* (Fox, 2006) and *openMX* () are available for download.

However R is entirely also entirely text and command-oriented. In order to specify a model such as that illustrated in Figure 1, the user must generate a model specification file with 74 lines of code as shown in Figure 2. Errors are easily made, and no additional tools are available to aid in the debugging process. Though the *openMX* and *sem* functions in R provide powerful analytic tools, the user interface does not appeal to researchers who feel their programming abilities are limited.

Proposed Project

The software required is an application that allows researchers to define their hypothesized models visually (as in Figure 1) and will output the correct syntax for the analytical software of their choosing (Figure 2).

To date a promising functional application has been developed in JAVA by a Computer Science student as a 4080 project. The existing software allows the user to draw a path diagram and outputs code for the R package sem. There are a number of improvements to be made (refinements and additions to graphical user interface) and then the application needs to be extended to output syntax appropriate for additional software applications (openMX, MPlus and EQS). Though this project may not begin at “the first stages” of the software lifecycle, this scenario is likely common in the software development market. In addition, the student will be working with a primary “client” who is far less technically advanced, which is also reflective of real-world situations.

The student is not expected to have any familiarity with statistics or the software packages mentioned above. Relevant details will be provided as needed.

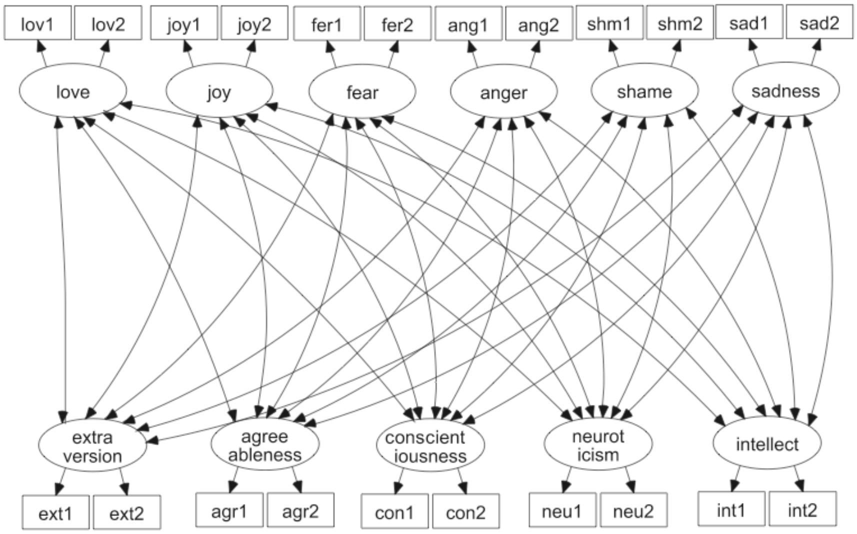
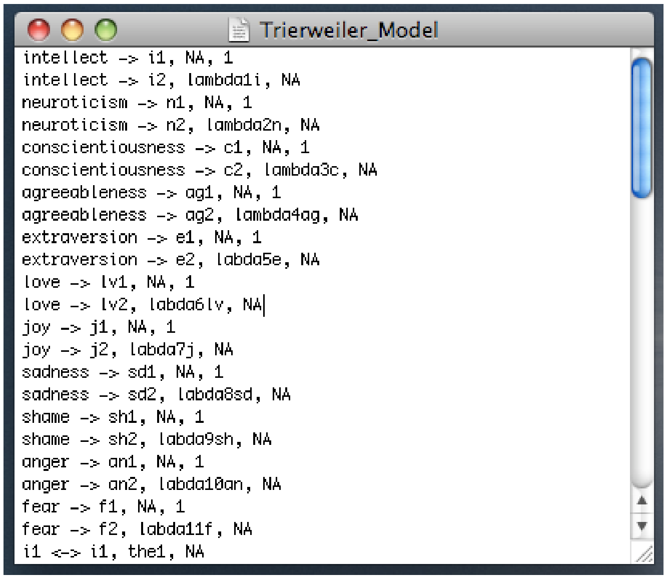


Figure 1: Path diagram of a theoretical model based on a study by Trierweiler, Eid, & Lischetzke (2002).

Trierweiler, L. I., Eid, M., & Lischetzke, T. (2002). The structure of emotional expressivity: Each emotion counts. *Journal of Personalaity and Social Psychology* , *82*, 1023-1040.



```
intellect -> i1, NA, 1
intellect -> i2, lambda1i, NA
neuroticism -> n1, NA, 1
neuroticism -> n2, lambda2n, NA
conscientiousness -> c1, NA, 1
conscientiousness -> c2, lambda3c, NA
agreeableness -> ag1, NA, 1
agreeableness -> ag2, lambda4ag, NA
extraversion -> e1, NA, 1
extraversion -> e2, labda5e, NA
love -> lv1, NA, 1
love -> lv2, labda6lv, NA
joy -> j1, NA, 1
joy -> j2, labda7j, NA
sadness -> sd1, NA, 1
sadness -> sd2, labda8sd, NA
shame -> sh1, NA, 1
shame -> sh2, labda9sh, NA
anger -> an1, NA, 1
anger -> an2, labda10an, NA
fear -> f1, NA, 1
fear -> f2, labda11f, NA
i1 <-> i1, the1, NA
```

Figure 2: Excerpt of model specification text file for the model based on a study by Trierweiler, Eid, & Lischetzke (2002).