Considerations When Evaluating the Accuracy of NGOMSL Models
A.Hornof - 5/11/01

• You should have carefully read Kieras’s NGOMSL Guide and have it close at hand when evaluating the models.

• Task instances should be sufficiently detailed so that an experienced user could do them from the description, and so that the task instances can drive the entire NGOMSL execution. For example, “Enter an appointment” is insufficient but ‘Enter ‘Dentist’ from 1 PM to 2 PM on May 15th” is sufficient.

• The NGOMSL model should be sufficiently detailed so that the model uses all of the details in the task instances. For example, methods for entering an appointment should at some point have the user enter the textual description of the appointment.

• Step through the methods as if you were a computer code interpreter and see if the models actually execute.

• Look for places where the model does not go all the way to the KLM level. For example:

  Method for goal: Change Text
  Step 1. Click in the text field.
  Step 2. Edit text field. // Bad! This is not an operator!
  Step 3. Return with goal accomplished

  In this method, Step 2 should be “Accomplish Goal: Edit text field.”

• Look for missing operators. For example, every Click should be preceded by a Point, and most Points should be preceded by a Locate. It is likely that operators will be missing for common subtasks such as entering text into a field on a web page. There are many steps involved. There are also many steps required to cut and paste a string of text—even if shortcut keys are used, there are actually three Locates and Points required—one Locate and Point to find the start of the text string, one to find the end of the text string, and one to find the new location of the text string.

• Look for cases in which the model is written such that the user would be forced to go through steps that he or she obviously would not execute for all task instances. For example:

  MFG: Create address book entry.
  Step 1. AG: Enter name.
  Step 2. AG: Enter address.
  Step 3. AG: Enter phone number.
  Step 4. AG: Enter email address.
  ...
  Step n. RGA

  This method assumes that, in every task instance, all of the information will be entered. Most of these AGs should be preceded by a Decide. For example:

  Step 2. AG: Decide: If address is given, Then AG: Enter address.

• Look for methods that are identical. Such methods should probably be a single more-generic method.
• When checking the learning and execution times, refer to Section 2.6 in the NGOMSL guide to remember how to count the various statements.

• For Ms, see the end of Section 2.2. And remember:
  – A Decide is not an M.
  – A Verify is an M.
  – Whether Recall_from_LTM or Retreive_from_LTM is an M is a judgment call, so the analyst should probably state his or her assumption for each. Note that GLEAN (a GOMS interpreter that runs on a computer and automatically always computes execution times) counts it as an M, but that once the information has been stored in working memory, no further retrieval time is required.

• You do not need to verify learning and execution times calculated by GLEAN.

Let me know what else you find and I should add to this document......