Using Linked Elements for creating Item Models of Multiple Languages

Karen Fung
Centre for Research in Applied Measurement and Evaluation
University of Alberta

Dr. Hollis Lai
Undergraduate Medical Education
University of Alberta

Dr. Mark J. Gierl
Centre for Research in Applied Measurement and Evaluation
University of Alberta

Paper Presented at the Annual Meeting of the Canadian Society for Studies of Education

Victoria, BC

June 4, 2013
Abstract

A new method known as linked elements for creating test items in multiple languages simultaneously using item models is presented here. Linked elements, alongside with n-layer or embedded elements, facilitate the use of Automatic Item Generation to increase its efficiency when multiple languages of a test are required. This process connects variables horizontally in an element to create instances that belong to the same languages, which shortens item generation time and solves the problem of complexities in an item model. An example in a medical examination context is shown in this paper for creating items in both English and Chinese. We further demonstrate the power of n-layer model and linked elements by generating items in English, Chinese, and Spanish simultaneously using a single item model.
Using Linked Elements for creating Item Models of Multiple Languages

Introduction

The traditional approach of test items translation is often time consuming and costly, as it requires translators who are proficient in two or more languages to perform translations item-by-item, usually from English to another target language. This process becomes more effortful when a large item bank is required, especially in the case of computer based testing (CBT) and computer adaptive testing (CAT) where large amount of items are needed in order to maintain a certain level of item exposure. With the vast growing of technology, Automatic Item Generation (AIG; Gierl & Haladyna, 2013) provides a feasible solution to allow many items to be generated in a short period of time, using computer technology. And recently, with the translation done at the item model level, items of multiple languages were generated simultaneously (Fung, Gierl, & Lai, 2013; Gierl, Fung, Lai, & Zhang, 2013). Translations of item models were performed using n-layer modeling which uses a “branching-off” hierarchical approach to generate items when a secondary element can be embedded into a primary element (Gierl & Lai, 2012). In comparison to its simpler 1-layer model, an n-layer model allows greater generation capacity. One problem arise, however when the number of elements being embedded increase, causing longer iteration time during the automatic item generation process. This is also the case for the translation of item models because language variables are often embedded elements in an n-layer model. In this paper, a new method known as linked element is introduced for use alongside with n-layer modeling to increase efficiency of AIG. Examples and detailed explanations will be provided for greater understanding of this approach in item modeling for the context of item model translations.

Automatic Item Generation

As the implementation of technology in test development grows, the emergence of Automatic Item Generation (AIG) is becoming more prominent. AIG is the use of computer technology to generate items automatically by using item models (Gierl & Haladyna, 2013; Gierl, Fung, Lai, & Zhang, 2013). It
addresses the problems with the traditional approach of item development which content specialists were needed to hand craft each item one-by-one, which can be very timely and expensive. With the use of computer programs, AIG make use of the item models with varying elements to create many different instances of item in a short period of time.

**Item Generation Process**

Recently Gierl, Lai, and Turner (2012) have described a three-step AIG process for the development of multiple-choice items in the medical context. In the first step, a *cognitive model for AIG* is developed with the expertise from content specialists. This cognitive model is a graphical representation in mapping the thinking processes for solving a particular medical problem. In order to develop the cognitive model for AIG, important key information required for solving a problem is identified. Such information is necessary for the next stage in item model development. An example of a cognitive model on the topic of solving a Hernia problem of surgery is presented in Appendix A. There are three layers to a cognitive model. The *Problems and Scenarios* as the first layer identifies the different scenarios for the topic of Hernia, and each of these scenarios has their associated *Sources of Information and Features* which leads to a different solution or conclusion of the problem. It is also the features that are brought forward to the item model to allow variations based on the scenario being created.

In the second step of AIG process, an item model is developed for the medical problem by content specialists. An item model is a representation of the features and the structures for the items to be generated (Bejar, 2002; Drasgow, Luecht, & Bennett, 2006; Lai, Alves, Zhou, & Gierl, 2009). It is created based on the parent item created by content expert, and the modification of the parent item is based on several components: the stem, options, and auxiliary information. The *stem* is the question statement that examinees will be answering; the *options* (in a multiple choice scenario) would be the choices given for examinees to choose their response for that item; the *auxiliary information* are
illustrations of figures, tables, and graphs that would appear on the stem, and/or in the options (Gierl et al., 2008; Lai et al., 2009). Altering these components using elements would provide a variety of generated items using different combinations. Elements are the manipulating variables to create different instances of the stems and the options. The use of constraints control the elements found in the final items generated. In the final step, the item model is fed into a computer program for item generation. In this step, all the components in an item models are combined together to create different instances of items through iterations of the varying elements. An example of an AIG computer program is a java based program by Gierl et al. (2008) known as IGOR (i.e., Item GeneratOR).

Types of Item Models and Linked Elements

There are two types of item models: 1-layer model and n-layer model. A 1-layer model is used when only a small number of elements are modified, and that instances of the generated items contain a smaller variability thus lesser amount of items being generated. It calls for a linear approach in the generation process, where each specified value in an element is matched with others horizontally. On the other hand, n-layer model allows for greater expansion of an item model, with more variability in item instances. In an n-layer model (also known as embedded elements), values of an element may have another element embedded within which also calls for other different values. The n-layer model uses a hierarchical (or a branching off) approach in generating more instances of an item. The use of an n-layer model allows an item model to have more generative power. A comparison of a 1-layer model and an n-layer model is presented in Appendix B. Figure 2 and Figure 3 show an example of the two types of item models for a medical licensure exam context. The difference can be found in the section Item Model - Stem where the 1-layer approach with the stem written out containing varying elements. However with the n-layer approach, the element Situation is stated where different combinations of scenario can be shown as the stem. Figure 4 provides an illustration in comparing the two types of item models. It can
be seen that an n-layer model allows greater generation capacity when elements are embedded within one another.

Multilingualism in AIG and the use of Linked Elements

Our previous research in AIG provided examples of how AIG can be used in different context (i.e., Mathematics, Science, and Medicine). And recently we were able to increase the efficiency of AIG by generating items of different languages (Spanish, and Chinese) in simultaneously with the original English items. This multiple-language approach of AIG uses n-layer modeling with language variables as embedded elements, so during the generation process variables of both languages will be called for at the same time. And the ending results would be a large number of items in two or more languages can be generated at the same time with a significant shorter period of time in comparison to the traditional approach of test translation where items have to be translated one at a time.

Translation of Item Models

As item models are the main resource for AIG, in order for items to be generated in multiple languages, the translation process needs to be performed at the item model level. A translation procedure of item models in the medical context was described by Fung, Gierl, and Lai (2013). There are two stages of the translation processes (i.e., non-expert translation and expert translation) and each stage involves three-steps (i.e., word-level translation, sentence-level translation, and item-level translation). In the first stage, a non-expert will translate the item model into the target language using the original English item model. In the first step, varying elements, terminologies, and individual words are first translated, including any technical and common terminologies. Then in the second step where the translated words from the first steps were combined into a sentence, this stage ensures the flow, grammar, and the structures of the sentences are proper. And in the final step of this stage, all sentences are combined together as a whole, in this step, all instances of elements will be taken into consideration of and to ensure the generated items will flow well structurally and grammatically in all
aspects. The purpose of having a non-expert to translate the model in the first stage is to save time for
the content expert to perform any translations based on the translated model by non-expert, providing
a more cost saving solution.

In the second stage, the translated item model by the non-expert will be shown to the content
expert, alongside with the generated items (both English and the target language) for comments and
suggestions on improving the translation. At this stage the content expert performs a quality check of
the translation done by the non-exert. This stage also ensures that the quality of the translated model
will fit the testing context in terms of how phrases should be structured and terminologies are used
appropriately. A similar three-step process was used in this stage, with varying elements, and medical
terminologies were being verified first, and then onto the sentence level to ensure the phrasing of
sentences are appropriate for the medical context. And finally verifying the item model as a whole to
ensure all instance of the items are considered. Any modifications of the translated item model would
be made from the non-expert translation unless major discrepancy occurs which the suggestion from
the content specialists will be used.

Linked Elements

One problem arises when language is added as an extra layer of component in an item model.
As more elements are embedded in an n-layer model, more “branching off” iterations are required
during the generation process. As a result of this additional layer, it will take a longer generation time
and the requirement of more powerful computers. Linked elements, in combination with the use of an
n-layer model, allows for greater efficiency in the generation process. Instead of creating different
variables that are expected to belong in the same category (e.g. language), linked elements allows for
grouping of such related variables. Each value in an element is matched in parallel with the values of its
associated element(s). This feature saves time for programming constraints in an item generator and
speeds up the generation process when the generator has lesser individual variables to go through.
Figure 4 illustrates the concept of linked elements in the context of generating items in English and Chinese together. As seen in the figure, before the elements were linked, they are seen as independent from one another in which constraint programming is required to provide instruction that English Element A and English Element B are related to each other. And once the linked element approach is used in the IGOR context, elements of the same languages are grouped together (with the assigning prefix “Eng.” and “Chin.”). During the iteration process, IGOR will automatically link all the elements with the same prefix together in a horizontal manner, thus preventing the need for constraint programming.

The benefit of linked elements over constraint programming is the difference in how a computer iteration program like IGOR reads the item model. With the use of constraints, it provides a restriction to the computer program on what should and should not be shown in the final items, after all the iterations have already been performed. However, with the use of linked elements, during the iteration process, when a prefix is read, IGOR would automatically combine the groups together while linking all their associated values horizontally and produce the items, without needing to iterate all possible combinations of items. In item model translation where elements of the same language are linked together, linked elements can function in four different forms: words, key phrases, single sentences, and multiple sentences (Gierl, Fung, Lai, & Zhang, 2013). These four forms are then used to adapt words, phrases, and sentences from one language to another to permit multilingual AIG. This study demonstrates the use of linked elements to group language as a category for generating items of multiple languages simultaneously.

**Purpose of Research**

The purpose of this research is to introduce an efficient method of using item models in the IGOR context to automatically generate test items in multiple languages simultaneously. Using a java based program developed by Gierl, Zhou, and Alves (2008) known as IGOR (Item GeneraTOR) with the
support of multiple languages, linked elements joins multiple groups that belong to the same language in a horizontal manner to create difference instances in the automatic item generation (AIG) process. Using examples and demonstration throughout the paper, we will demonstrate the usefulness and effectiveness of the linked elements approach.

**Methodology**

An item model developed in previous studies for a medical licensure exam was chosen for this study. The item model is in n-layer, in English, and on the topic of Hernia repair in the context of surgery. In this study, we will demonstrate the use of linked elements in creating items of multiple languages simultaneously (i.e., English, Chinese, and Spanish). Translation of the item model was performed using the 2-stage (non-expert translation and expert translation) method described previously. For the first part of the study, linked element was used to create an English and Chinese item model. The non-expert was the primary researcher who is fluent in both English and Chinese and had no prior education in medicine. The non-expert first translated the original English item model into Traditional Chinese using the three-step processes, from word and terminology level, to the sentence level and then further move onto the item block level. Upon completion of the translation of item model, both the Chinese and the English versions were combined into a single model and items of both languages were being generated to provide a basis for the content expert to make any further changes on the quality of translation. The content expert is an experienced surgeon who is proficiency in both English and Chinese. The task of the content expert was to provide improvement suggestions on the translation so the items generated would appropriately fit with the medical context. A similar 3-step process was used to make verifications on the translation.

To use linked elements for combining the two languages together into a single model, elements of the same language were linked as the same category, so during the generation process using IGOR,
when one element of that linked is being called, its associated elements will also be called. In this proposal, an item from the medical licensure exam in surgery is shown.

**Results and Conclusions**

In this example of the surgery item that requires diagnosing complications associated with hernia, a total of 360 (180 each in English and Chinese) were generated. A sample of these items can be found in Appendix C. Upon doing manual check by the researchers who are fluent in both English and Chinese and a Chinese speaker, the outcome was as expected. All the linked elements came out as expected and both languages were being generated simultaneously in a very quick manner. This finding allows us to conclude that the linked element is a useful tool for grouping and generating items in different languages at the same time. We are now in the process of combining the Spanish and the Chinese version of the Hernia model into the English model so three languages can be generated at the same time. In addition to the science item models we used, we are also expanding the use of linked elements in other exams to increase its generalization.

**Educational Importance of Study**

This study demonstrated the use of linked elements, along with n-layer models to generate items of different languages simultaneously in a short period of time. The results can benefit the field of test development and further benefit students of different languages. It saves time and resources when one item model can be used for more than just one student population, to which the saved resources can be allocated into something else such as development of curriculum.
References


Appendix A

Figure 1. Cognitive model for AIG using hernia example.

Structure of content knowledge for issues related to Hernia

**Features**

- **Acuity of Onset**
  - Elements
    - Range of Time
    - Constraint
      - As: Months/Years
      - P: /-6 hours over days
      - M: More than 5 hours
      - A: Any time

- **Pain**
  - Element
    - No pain
    - Constraint
      - P: Intense-severe
      - S: Severe
      - M: Moderate-mild

- **Nausea and Vomiting**
  - Element
    - Not nauseous and vomiting
    - Constraint
      - P: Much
      - S: Very body

- **Groat Pain**
  - Element
    - Groat pain present
    - Constraint
      - Any one of the element will present in all scenarios

- **Umiliscus Pain**
  - Element
    - Umiliscus pain present
    - Constraint
      - Any one of the element will present in all scenarios

- **Scars**
  - Element
    - Scar presented from previous surgery
    - Constraint
      - Any one of the element will present in all scenarios

**Sources of Information**

- **Patient Presentation**
- **Location**
- **Physical Examination**
- **Laboratory Results**

**Problem and Scenarios**

- **Asymptomatic Incarcerated (AI)**
- **Painful Incarceration (PI)**
- **Strangulation (S)**
- **Reducible Symptomatic (RS)**

**Hernia**

**White Blood Count**

- **Element**
  - Normal
    - Constraint
      - E: Increase
**Figure 2.** One-layer model in the context of surgery.

<table>
<thead>
<tr>
<th>Parent Item:</th>
</tr>
</thead>
<tbody>
<tr>
<td>A 24-year-old man presented with a mass in his left groin. It appeared suddenly 2 hours ago while lifting a piano. On examination he has a tender firm mass in the left groin. Which one of the following is the next best step?</td>
</tr>
<tr>
<td>(A) Immediate hernia repair</td>
</tr>
<tr>
<td>(B) Needle aspiration</td>
</tr>
<tr>
<td>(C) Ice packs to groin</td>
</tr>
<tr>
<td>(D) Reduction of mass</td>
</tr>
<tr>
<td>(E) Ultrasound of groin</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Item Model:</th>
</tr>
</thead>
<tbody>
<tr>
<td>A [AGE]-year-old [GENDER] presented with a mass [PAIN] in [LOCATION]. It occurred [ACUITYOFONSET]. On examination, the mass is [PHYSICALFINDINGS] and lab work came back with [WBC]. Which of the following is the next best step?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Elements</th>
</tr>
</thead>
<tbody>
<tr>
<td>[AGE] (Integer): From 25.0 to 60.0, by 5.0</td>
</tr>
<tr>
<td>[GENDER] (String): 1: man 2: woman</td>
</tr>
<tr>
<td>[PAIN] (String): 1: 2: and intense pain 3: and severe pain 4: and mild pain</td>
</tr>
<tr>
<td>[LOCATION] (String): 1: the left groin 2: right groin 3: the umbilicus 4: an area near a recent surgery</td>
</tr>
<tr>
<td>[ACUITYOFONSET] (String): 1: a few months ago 2: a few hours ago 3: a few days ago 4: a few days ago after moving a piano</td>
</tr>
<tr>
<td>[PHYSICALFINDINGS] (String): 1: protruding but with no pain 2: tender 3: tender and exhibiting redness 4: tender and reducible</td>
</tr>
<tr>
<td>[WBC] (String): 1: normal results 2: normal results 3: elevated white blood cell count 4: normal results</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Options</th>
</tr>
</thead>
<tbody>
<tr>
<td>exploratory surgery; reduction of mass; hernia repair; ice applied to mass</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Key</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conditional</td>
</tr>
</tbody>
</table>
**Figure 3.** *N-layer model in the context of surgery.*

<table>
<thead>
<tr>
<th>Item Model:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stem</strong></td>
<td>[Situation TestFindings QuestionPrompt]</td>
</tr>
</tbody>
</table>

**Elements:**

**QuestionPrompt (Text):** 1. What is the best next step? 2. Which one of the following is the best prognosis? 3. Given this information, what is the best course of action?

**TestFindings (Text):** 1. On examination, the mass is [[PhysicalFindings]] and lab work came back with [[WBC]]. 2. Upon further examination, the patient had [[WBC]] and the mass is [[PhysicalFindings]]. 3. With [[WBC]] and [[PhysicalFindings]] in the area, the patient is otherwise nominal. 4. There is [[PhysicalFindings]] in the [[Location]] and the patient had [[WBC]].

**Layer 1**

**Situation (Text):** 1. A [[AGE]]-year-old [[Gender]] presented with a mass [[Pain]] in [[Location]]. It occurred [[AcuityofOnset]]. 2. Patient presents with a mass [[Pain]] in [[Location]] from [[AcuityofOnset]]. The patient is a [[AGE]]-year-old [[Gender]]. 3. Patient complaints of a mass [[Pain]] in [[Location]] which has been a problem since [[AcuityofOnset]]. 4. A [[Gender]] was admitted with pain in the [[Location]] from [[AcuityofOnset]].

- **[AGE]** (Integer): From 25.0 to 60.0, by 5.0
- **[GENDER]** (String): 1: man 2: woman
- **[PAIN]** (String): 1: 2: and intense pain 3: and severe pain 4: and mild pain
- **[LOCATION]** (String): 1: the left groin 2: right groin 3: the umbilicus 4: an area near a recent surgery
- **[ACUITYOFONSET]** (String): 1: a few months ago 2: a few hours ago 3: a few days ago 4: a few days ago after moving a piano
- **[PHYSICALFINDINGS]** (String): 1: protruding but with no pain 2: tender 3: tender and exhibiting redness 4: tender and reducible
- **[WBC]** (String): 1: normal results 2: normal results 3: elevated white blood cell count 4: normal results

**Options**

exploratory surgery; reduction of mass; hernia repair; ice applied to mass

**Key**

Conditional
Figure 4. A comparison between 1-layer modeling and n-layer modeling.
Figure 5. The concept of linked elements.

Linked elements in AIG 16

Individual elements in English and Chinese

Linked elements by language
Appendix C

Generated items for a sample item model (options are not shown in this example) – English

21. A 25-year-old woman presented with a mass in the left groin. It occurred a few months ago. On examination, the mass is protruding but with no pain and lab work came back with normal vitals. What is the best next step?

25. Patient presents with a mass in the left groin from a few months ago. The patient is a 25-year-old woman. Upon further examination, the patient had normal vitals and the mass is protruding but with no pain. What is the best next step?

29. A woman was admitted with pain in the left groin from a few months ago. There is protruding but with no pain in the left groin and the patient had normal vitals. What is the best next step?

6. Patient presents with a mass and intense pain in the right groin from a few hours ago. The patient is a 25-year-old man. Upon further examination, the patient had normal vitals and the mass is tenderness. What is the best next step?

10. Patient complaints of a mass and intense pain in the right groin which has been a problem since a few hours ago. With normal vitals and tenderness in the area, the patient is otherwise nominal. What is the best next step?

14. A man was admitted with pain in the right groin from a few hours ago. There is tenderness in the right groin and the patient had normal vitals. What is the best next step?

41. Patient presents with a mass and intense pain in the right groin from a few hours ago. The patient is a 40-year-old man. Upon further examination, the patient had normal vitals and the mass is tenderness. What is the best next step?

37. A 40-year-old man presented with a mass and intense pain in the right groin. It occurred a few hours ago. On examination, the mass is tenderness and lab work came back with normal vitals. What is the best next step?

47. A 40-year-old woman presented with a mass and intense pain in the right groin. It occurred a few hours ago. On examination, the mass is tenderness and lab work came back with normal vitals. What is the best next step?

7. Patient presents with a mass and severe pain in the umbilicus from a few days ago. The patient is a 25-year-old man. Upon further examination, the patient had elevated white blood cell count and the mass is tender and exhibiting redness. What is the best next step?
85. 一名 40 歲的女病人在右側腹股溝出現一團帶有強烈痛楚的組織。病徵已延續了幾個小時。經檢查後，那組織是軟的，化驗結果顯示命脈正常。那最佳的下一部是應該怎樣？

92. 一名病人的腫脹從數天前出現一團帶有腫痛的組織。病人年齡男，25 歲。經長細檢查後，病人白血球上升，而那組織是帶紅而軟的。那最佳的下一部是應該怎樣？

93. 一名病人的腫脹從數天前出現一團帶有腫痛的組織。病人年齡女，25 歲。經長細檢查後，病人白血球上升，而那組織是帶紅而軟的。那最佳的下一部是應該怎樣？

94. 一名女子因數天前腫脹出現腫痛。在腫脹上感覺到帶紅而軟的，而病人的白血球上升。那最佳的下一部是應該怎樣？

95. 一名 25 歲的女病人在腫脹出現一團帶有腫痛的組織。病徵已延續了數天。經檢查後，那組織是帶紅而軟的，化驗結果顯示白血球上升。那最佳的下一部是應該怎樣？

96. 一名病人的腫脹從數天前出現一團帶有腫痛的組織。病人年齡女，25 歲。經長細檢查後，病人白血球上升，而那組織是帶紅而軟的。那最佳的下一部是應該怎樣？

97. 一名女子因數天前腫脹出現腫痛。在腫脹上感覺到帶紅而軟的，而病人的白血球上升。那最佳的下一部是應該怎樣？

107. 一名病人手術的傷口附近從數天前，自剛剛搬移鋼琴後出現一團帶有微痛的組織。病人年齡男，25 歲。經長細檢查後，病人命脈正常，而那組織是軟而可縮小的。那最佳的下一部是應該怎樣？

108. 一名病人正憂慮數天前，自剛剛搬移鋼琴後在手術的傷口附近出現的一團帶有微痛的組織。病人命脈正常，加上受影響範圍軟而可縮小的。病人大概一切正常。那最佳的下一部是應該怎樣？

109. 一名男子因數天前，自剛剛搬移鋼琴後手術的傷口附近出現腫痛。在手術的傷口附近上感覺到軟而可縮小的，而病人的命脈正常。那最佳的下一部是應該怎樣？