COMPUTER MATH SNAPSHOTS - COLUMN EDITOR: URI WILENSKY*

Agents with Attitude: Exploring Coombs Unfolding Technique with Agent-Based Models

Michelle Hoda Wilkerson

1 Introduction

This snapshot is intended as an illustration of some ways that computational modeling can help one engage with and extend pragmatically complex applied mathematical problems. Specifically, I explore Coombs unfolding technique (1950), a mathematical method used in psychometrics to analyze individual and collective preferences toward some topic of interest. Coombs unfolding works well when used on individuals, but becomes difficult if one wishes to combine individual data to determine the attitudes and preferences of an entire population. In this paper, I briefly introduce Coombs unfolding technique, and model the technique using simple agent-based models (ABMs) developed in the NetLogo (Wilensky 1999) modeling environment. I then attempt to address some of the difficulties one may encounter when applying Coombs unfolding to a population by using these ABMs as a starting point. Finally, I discuss the affordances and disadvantages of using computational modeling as a starting point for thinking about applied mathematical problems.

2 Coombs Unfolding: An Overview

Imagine you are asked to complete a survey that is designed to measure your position regarding topic U, where U may represent, for example, level of liberalism. Often, such surveys ask you to select from a list of items that represent differential levels (higher or

---

M. H. Wilkerson (✉)
Learning Sciences Program, Northwestern University, 2120 Campus Drive, Evanston, IL, USA
e-mail: m-wilkerson@northwestern.edu

lower) of topic U which you feel is closest to your own opinion regarding the topic. In the case of liberalism, these items may take the form of specific politicians. For analysis, these items are typically linearly mapped to numerical values, such that increasingly liberal politicians W, X, Y and Z take on values 0, 1, 2, and 3 on the scale of liberalism. There are several assumptions implicit in the use of items to measure an individuals underlying preference regarding some topic of interest:

1. *The topic under investigation is unidimensional*. Participants are not asked to rank politicians simply on their preferences for each; rather they are asked to do so specifically in terms of the topic of interest, liberalism.
2. *Individuals can only have one preferred location along the scale at any time*. A participant cannot report that they prefer both politicians X and Z with regard to liberalism, because it implies that they simultaneously prefer a moderately low as well as a very high level of liberalism.
3. *Items must be ordered*. In order for our survey to work well, respondents must agree that the level of liberalism of each available politician increases from W to X, from X to Y, and so forth. The items are not reliable if two respondents both prefer a very high level of liberalism, but one respondent believes that politician X exhibits this highest level and the other believes politician Z does.
4. *Items must be equally spaced from one another with regard to the topic of interest*. In order for politicians W, X, Y and Z to be mapped to values 0, 1, 2 and 3, respondents must believe that if politician X is n more liberal than politician W, then politician Y must similarly be n more liberal than politician X and so forth.

Certainly it is difficult to imagine topics and items for which all four of these criteria apply. It becomes much more conceivable to do if we get rid, however, of the fourth criterion: that items must be equidistant along the scale. This assumption has several consequences: it means that if an individual selects any given item, say Y, they are equally close, psychologically speaking, to items X and Z; that is, an individual who selects politician Y is equally likely to select X or Z if Y is no longer available. It also means that all individuals share the same perceived distance between items along the topic of interest. It seems, however, much easier for different individuals to agree on the order of items along a scale than distance in this way.

In 1950, Clyde Coombs introduced a method that eliminates this problematic fourth criterion from survey questions while preserving the first three, enabling such survey questions to apply appropriately to a greater number of phenomena. Unfolding technique is a method for measuring individual's preferences toward some topic U without making unwarranted assumptions about the relative distance of the items used to measure U. Rather than ask respondents to identify what item from a set of options they think is best, Coombs proposed that they rank-order items from highest to lowest in preference, expectation, or endorsement. Using this ranking, one can figure out whether a given individual thinks certain items are closer to one another with respect to the topic of interest. Hence, given a respondents ranking of politicians W, X, Y and Z (where each politician's unfolding can help determine whether that respondent perceives X and Y to be closer in their level of liberalism than W and X, or Y and Z.

To better understand unfolding technique, it makes sense to walk through the technique with a few specific examples. Assume, as discussed above, that a survey respondent is asked: Please rank politicians W, X, Y, Z by how closely they represent your level of liberalism. Respondents R_1, R_2, and R_3 provide three different rank-orderings I_n: $I_1 = YZXW$, $I_2 = YXWZ$, and $I_3 = YXZW$. Note that in all three cases, respondents
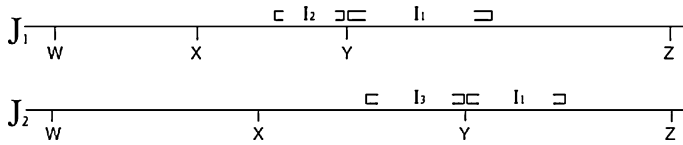
Fig. 1 Respondents R_1 and R_2 each must be located in intervals I_1 and I_2, respectively, along J_1, since these intervals each allow the respondents to be progressively "further" from the politicians they prefer less. Respondent R_3's rank-ordering, however, is not compatible with J_1 because there is no spatial interval for which politicians Y, X, and Z are all closer than politician W. At the same time, both R_1 and R_3 are compatible with J_2, but R_2 is not

identified politician Y as the most preferable: that is, all three respondents are closest along the liberalism scale to politician Y. Their perceptions of how close each politician is to one another, however, are very different. Respondents who report a rank order I_2, for example, can only rank-order their preferences as Y, X, W then Z if they believe politicians X and W are both closer in liberalism to Y than Z is to Y—a view of the liberalism scale that cannot be accommodated by traditional interval interpretations of successive items along that scale. Respondents who report a rank order I_3, however, cannot consider herself to be closer to Y, X, and Z than W unless she understands W to be more distant from Y in terms of liberalism than Z. This difference, implicit in rank orderings I_2 versus I_3, can be illustrated using space as a metaphor for psychological distance (Fig. 1). Notice that, in this case, ranking I_1 can be accommodated by either of the two arrangements of politicians (1a and 1b). Each different arrangement is called a joint scale, or J_m.
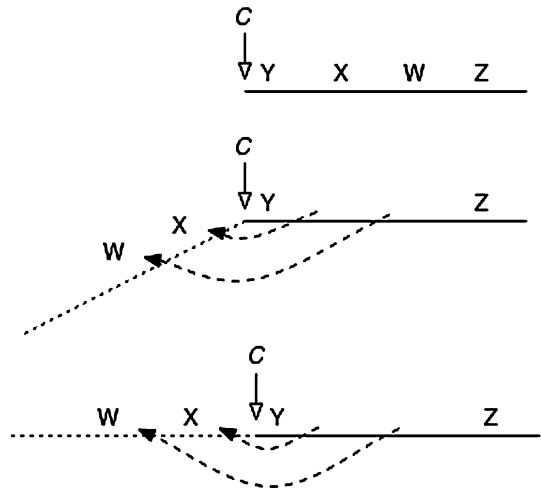
The technique for creating joint scales is called unfolding because spatially, an individual's reported rank-ordering of items I_n, can be unfolded like a piece of paper to reveal a joint scale that accommodates the reporting individual's perceived relative psychological distance for each item with regard to the topic of interest. The reporting individual's preferential position along that scale, C, is identified as the location along which the scale is folded when the rank ordering is reported (Fig. 2).

Unfolding becomes especially useful but also especially difficult when it is used to determine trends in the attitudes and preferences of a population, rather than of a single individual. That is because while several individuals reported rank orderings can often be accommodated by the same joint scale, other rankings are incompatible. Using this fact, and the Coombs unfolding technique in general, one can ask: Is there an explicitly agreed-upon, or commonly perceived, distribution of items along a dimension? What is the distribution of a given population's preferred locations along that measure? In the next section, I will begin to address these difficulties by using ABMs to explore unfolding in the context of a population. I will do so using Coombs own example of unfolding theory in the context of student's reports of expected grades in a psychology class (1950).

3 Agents with Attitude: Four Agent-Based Models for Exploring Coombs Unfolding

To better understand Coombs unfolding technique in the context of populations, I developed a series of ABMs that each emphasize a different aspect or application of unfolding theory. ABMs are particularly well suited for such an exploration, because they allow one to quickly and easily model large groups of individuals or agents (thus enabling us to

Fig. 2 A joint scale J_m is produced by "unfolding" a respondent's rank-order information. The respondent's own position along the scale is indicated by point C, and items are positioned to the left or right of the point C in the way that restores the predetermined order (W, X, Y, Z)

compare and combine those agent's preferences). NetLogo (Wilensky 1999), one such programming environment, also allows agents to live and move in a visual world called the view. This allows psychological distance to be modeled as physical distance in the model interface, and enables users interacting with the models to see a representation of a joint scale J_m, each agent's preferred location along that scale C, and the aggregate agent group's overall preferences with regard to the topic of interest (that is, collections of agents in each interval along the scale). The following models investigate four aspects of unfolding:

- *Model 1: Unfolding* automates the process of constructing one instance of a joint scale that will accommodate a given rank ordering of grades.
- *Model 2: Problematizing the Interval Scale* exhibits why a scale that evenly distributes items along one dimension does not accommodate all possible rank-orderings of grades.
- *Model 3: Deriving a Mutual Joint Scale* determines, given a population whose individual rank orderings are derived from a mutually agreed upon joint scale, what that joint scale is.
- *Model 4: Negotiating a Joint Scale of Best Fit* offers one method of compromising between conflicting joint scales when using rank order data to determine population-level attitudes.

To provide some context for the models, I will use an example that was elaborated by Coombs himself in his paper introducing unfolding technique: students' expected grades in a course. Expected grades satisfy the conditions enumerated above fairly well: they are unidimensional (based on ones performance in class), students can only expect one grade at a time (that is, they cannot expect to simultaneously receive a B and a D in the class), and it is unlikely that the order of grades is under contention (that is, there are not cases where a lower grade reflects better performance than a higher grade).

In this example, each student, R_n, provides a rank-ordering of grades (I_n). As described above, this ordering limits the number of qualitatively different scales (J_m) for which the student's preferred position (C_n) will be accommodated. In this example, different J_m scales represent differences in the assumed psychological distance between different

grades. I explore cases in which the joint scale for a class cannot be assumed to be interval—that is, grades are not necessarily evenly distributed along the performance dimension. For example, the professor of this course may be well-known for awarding the grade A to only a select few students who produce exemplary work, creating a large psychological distance between A and other grades; or alternatively for never having failed a student.

Model 1: Unfolding. The first model in the Coombs series takes an ordered ranking of grades as input, and "unfolds" this ranking (see Fig. 2) to create a spatial arrangement of items that is guaranteed to accommodate the ranking provided. This model illustrates the simplicity with which unfolding can be accomplished on an individual basis, and also illustrates the difficulty one encounters when trying to combine these unfolded scales in order to make claims about the attitudes of a population.

```
Given a rank−ordered list of k items, I1 to Ik, call the nth item In.
Place I1 in the center of the screen.
Set n = 2.
While n < k,
   If In > I1, place it to the right of I1 n units away.
   If In < I1, place it to the left of I1 n units away.
   If n = 2, place the individual 1/4th of the distance from I1 to I2.
   Set n = n + 1.
```

While this technique does a good job of identifying and accommodating all psychological distance information provided in a given rank ordering, it does not help one determine which individual scales can coexist along the same joint scale. In fact, the rank ordering provided by $I_{4^{th} row}$, ABCDF, can be accommodated by all of the scales constructed by the unfolding model, since all that is required is for the individual to be closest to item A, next closest to item B, and so forth. To address this, the next model illustrates how any one scale can accommodate 8 out of 16 possible ordered rankings of grades (Fig. 3).

Model 2: Problematizing the Interval Scale Indeed, any joint scale can accommodate a number of (though not all) different rank orderings of grades. This model illustrates which rank-orderings of grades can and cannot be accommodated by an interval scale, in which there is equal space between each item. It does so by comparing each agent's rank-ordering to their relative distance to each item along the scale. If an individual finds a position along the interval scale at which their closest-to-furthest distances to each item matches their rank-ordering of closest-to-furthest items, they are considered to be accommodated by that scale.

Fig. 3 Grade scales determined using the unfolding model. Rank orderings inputted were $I_{1^{st} row} = BCADF$, $I_{2^{nd} row} = DFCBA$, $I_{3^{rd} row} = BCDAF$, $I_{4^{th} row} = ABCDF$

```
Create n agents.
Give each agent a random rank−ordering, I1 to Ik.
Make all of the agents unhappy.
Randomly place agents in the world.
Place items in the world with equal space between each item.
While there are any unhappy agents,
   Ask the unhappy agents to step randomly to the right or left.
   Ask agents to make a list of items, from nearest to furthest.
   If the agent's "nearest to furthest" list = the agent's rank ordering,
     Make that agent happy.
```

This model illustrates that some item distributions (such as an interval scale) can accommodate many individual rank orderings at one time, but not all. In the case of five items, there are 16 possible individual scales—that is, 16 different rank-orderings that preserve ordinal relationships. If we assume a five-item *interval* scale, eight of these individual scales are preserved (such as [0 then 1 then 2 then 3 then 4], the navy blue group in the screenshots above, or [3 then 4 then 2 then 1 then 0], the peach group). By simply changing the distances between items in the world, other scales can be tested using this same algorithm to determine which individual rankings are accommodated and which are not.

The algorithm, however, starts with an already-assumed joint scale. It does not allow a joint scale to be ascertained from given individual data; which is the primary goal of unfolding theory. The next two models I present begin to deal with the difficulty of using the rank-ordered grades provided by a number of individuals in a population to figure out whether all members of that population can be accommodated by a mutual joint scale (Fig. 4).

Model 3: Deriving a Mutual Joint Scale In order to begin to explore how one can use individual-level data in the form of rank orderings to find a more generally agreed-upon joint scale, it makes sense to first start with a simple case: one in which a population agrees on the arrangement of items along the scale of interest, and only differ in their own preferred or expected location along that scale. For example, we might imagine a course in which all students agree that the effort required to receive an A is much higher than that required to receive a B, but in which many students expect to receive different grades. The following algorithm uses spatial distance to determine an ideal position for items along a scale (that accommodates all individuals' rank orderings), and then locates each individual within that scale.

```
Randomly determine an assumed underlying scale, J.
Create n agents, each with a rank−ordering generated from J.
Randomly place agents in the world.
Randomly place items in the world.
Make agents unhappy.
Mark each collection of agents with identical rank−orderings as a group.
While there are any unhappy agents,
```
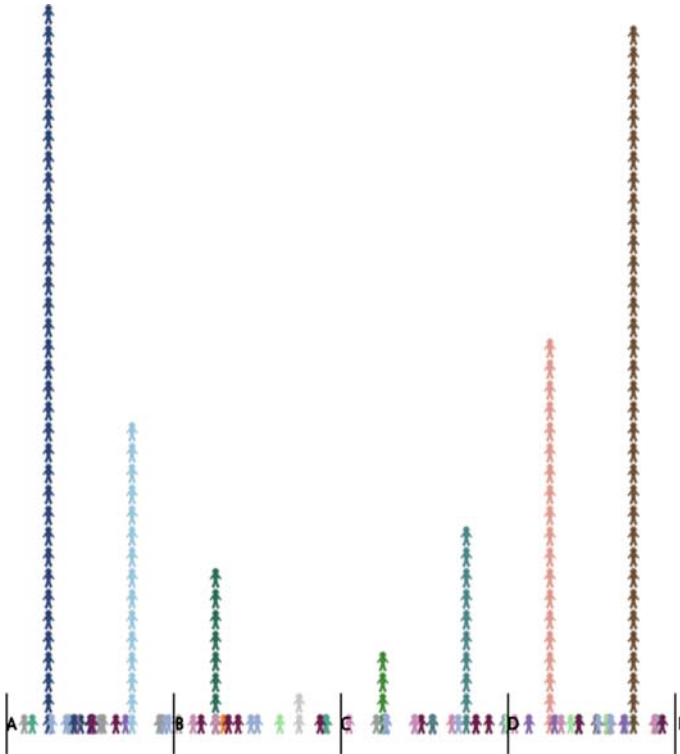
Fig. 4 If we assume an interval scale (items are located at the bottom of the world), only 8 out of 16 possible rank-ordering of grades are accommodated. Agents whom are accommodated by the interval scale comprise a histogram above the items, while agents along the base of the world are unhappy and still seeking a point at which their ordered rankings are accommodated

```
Ask the unhappy agents to step randomly to the right or left.
Ask agents to make a list of items, from nearest to furthest.
If the agent's "nearest to furthest" list = the agent's rank ordering,
    Make that agent happy.
If there is any group of agents with no happy people
    Ask one of the items to step randomly to the right or left.
```

This model does a good job of identifying the required relative distances between items that define a common underlying joint scale. It also does a good job of showing how different individual scales that are derived from the joint scale are distributed, and the preferences that they produce. However, because items move randomly until all individual scales are accommodated, some differences in distance are simply reflective of random movement, and not of actual psychological distances as manifested in agent data. Furthermore, all agents must share the same underlying scale in order for this method to be effective; it cannot determine more than one underlying scale within the population (Fig. 5).

Model 4: Negotiating a Joint Scale of Best Fit. How, then, might one begin to reconcile rank-ordering information that might not reflect a single global joint scale shared by all
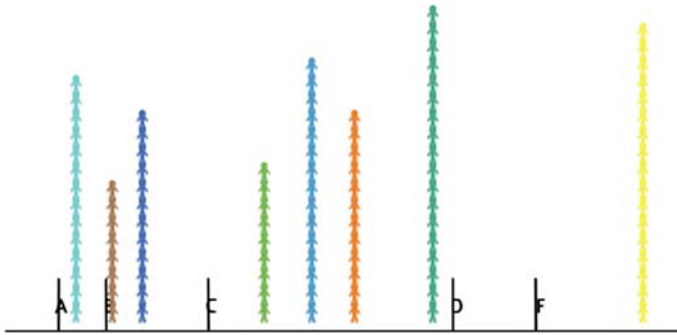
Fig. 5 In the perfect case, all of the rank-orderings that are reported by agents can be accommodated by the same joint scale. The various rank ordered grades reported by the population featured in the model are all reconcilable with a scale that assumes that the grades *A* and *B* are psychologically the closest, and *C* also relatively close, but that one must perform significantly worse in the course to receive a *D* rather than an *A*, *B*, or *C*

respondents, but that still tells us something about population-level opinions of some topic? To use our example, if we are given a set of rank-orderings of grades that are not all accommodated by a single joint scale, what can we say about the population of students, and why? Here, I propose one strategy that begins to address this problem by searching for the arrangement that minimizes the number of respondents that are not accommodated by a given arrangement of items along the joint scale. This algorithm, like the previous one, uses random movements of both values and individuals in space to determine a spatial joint scale. In addition, the algorithm keeps track of each arrangement's fitness, and periodically resets to the best-fitting arrangement if a better one cannot be found in a given number of random re-arrangements.

```
Create n agents, each with a random rank−ordering.
Randomly place agents in the world.
Randomly place items in the world.
Make agents unhappy.
Mark each collection of agents with identical rank−orderings as a group.
Set Count to 0.
While there are any unhappy agents,
   If Count > 500
     Ask the unhappy agents to step randomly to the right or left.
     Ask agents to make a list of items, from nearest to furthest.
       If the agent's ''nearest to furthest''list = the agent's rank ordering,
        Make that agent happy.
       If there is any group of agents with no happy people
        Ask one of the items to step randomly to the right or left.
     If the number of agents that are unhappy is lower than least,
       Record the current arrangement
```
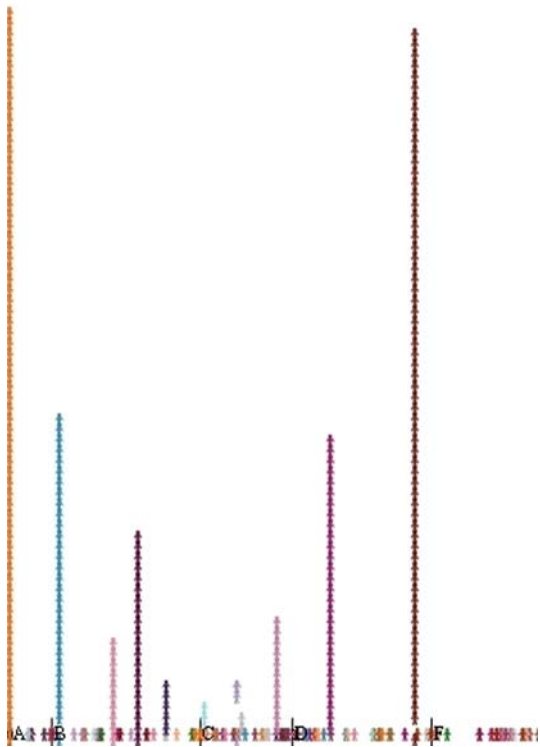
Fig. 6 A partial solution—given random setup conditions, this model produced a scale that accommodates 10 different rank-orderings of grades, many with a large population of individuals that report each rank order


```
        Make Least = number of unhappy agents
         Set Count to 0
      Else
         Set agents and values to last best arrangement
         Set Count to 0
```

Certainly this solution is not the most computationally efficient. It does, however, demonstrate how, in theory, more complex rank-order data can be negotiated and made meaningful using a strategy similar to the solutions I have proposed for simpler versions of the Coombs technique problem (Fig. 6).

4 Discussion

The Coombs technique, and especially the application of this technique to population-level data, is an interesting mathematical exercise. It is imprecise in that it deals with qualitative distances rather than precise measurements, and with individual data rather than aggregated statistical properties. This makes it difficult to quantify joint scales or individual locations along them—the very things one needs in order to construct population-level descriptions of the data. In this paper, I confronted the difficulties inherent in applying Coombs unfolding technique to a population of individuals by using ABMs. These models are especially appropriate for addressing such problems for several reasons:

–   AGENT-LEVEL COMPUTATIONS. Coombs technique is applied at the single agent level. Therefore, even if one wishes to make claims about a population more generally, they must perform calculations on each individual. ABM emphasizes the importance of computation at the agent level, especially in the context of large numbers of agents being involved in such computation at the same time.

–   EMERGENT AGGREGATE RESULTS. In the models above, and in the NetLogo modeling environment more generally, the population-level patterns that emerge from agent-level computations are easily visualized. In the case of Coombs unfolding, this visualization provides information about the relative distances of items along a joint scale of interest—the main result of interest.

–   RANDOMNESS. Agent-based models allow for the randomization of agent-level properties and processes such as an agent's rank-ordering, an agent's location along a joint scale, and the location of items along a joint scale. This randomization is essential for many of the solution strategies employed in the models featured above: for example, for finding an ideal arrangement of items given reconcilable data, and for finding an arrangement of best fit for irreconcilable data.

–   SIMPLICITY. In the models I created, simple (qualitative) mathematical rules were easily translated into simple computational rules that did not need traditional measures or quantification. Indeed, in all models, a given joint scale's accommodation of different rank orderings was entirely determined by the agent's *relative* distance to items along the scale, rather than by more specific or precise measures.

At the beginning of this paper, I suggested that computational modeling can help one to engage with and extend complex mathematical problems. Coombs unfolding, at its core, is not particularly difficult. Applying Coombs unfolding to a population, however, presents a number of problems. Agent-based modeling, a specific type of computational modeling that allows one to explore simultaneous individual-level interactions within a large population, allows Coombs unfolding to be performed on many agents at once, producing an *emergent* population level outcome. In fact, one might imagine that the algorithms that might have been developed to accommodate more precise, numerical data would have been much less transparent and more difficult than the models I include above. In this case, computation is not necessarily defined by the pursuit of precision. Instead, it is an exploration of just how far an idea can go when combined with powerful tools, difficult problems, and yes simple methods.

References

Coombs, C. (1950). Psychological scaling without a unit of measurement. *Psychological Review, 57*(3), 145–158.

Wilensky, U. (1999). NetLogo. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL. http://ccl.northwestern.edu/netlogo/