The Institute for Perception

For technical reports from The Institute for Perception or for information about short courses, please visit www.ifpress.com or email us at mail@ifpress.com

Reprinted from IFPress (2010) 13(1) 2,3

Optimum Product Selection for a Drivers of Liking® Project

John Ennis, Daniel Ennis and Charles Fayle

Background: A project team selecting products for a category appraisal must contend with a number of competing objectives. One of these objectives is to obtain information on the performance of its own products and those of its competitors. This usually leads to extensive discussion and debate among the product development and market research members regarding the inclusion of different products in a study with a limited number of product slots. A second objective, often overlooked and of more interest to the research analysts, is the need to span the drivers of liking sensory space so that a robust account of the resulting maps is obtainable. Such an account is essential if the model that underlies the landscape will be used to predict the acceptability of products not tested in the original study[1].

Products under consideration for inclusion in such an appraisal can be described on multiple attributes and there are numerous ways of considering the similarities and differences among the products. The attributes are often obtained from expert descriptive panels or from analytical data such as fat content, sucrose levels, or acidity. If there is existing knowledge of the relationship between these attributes and liking, then it is desirable to focus attention on those variables that are most likely to emerge as drivers of liking. For the purpose of this report, we assume that a set of likely drivers of liking has been chosen. Using this set we examine candidate products for consumer testing with the goal of finding a subset of products that spans the space of those variables as completely as possible. Although conventional solutions to this problem exist, including the inspection of principal component plots and cluster analysis of the products based on their attribute values, these methods are often descriptive and their interpretations are somewhat subjective. In this report we describe a novel method that provides definitive guidance by way of optimal solutions obtained using concepts from the mathematical field of graph theory.

Scenario: You are interested in developing a landscape map of carbonated orange-flavored beverages and you have eighteen candidate products under consideration for inclusion in the consumer testing component of your project. Budget and practical testing constraints require you to choose twelve products for testing and you are interested in finding the best twelve of the eighteen to choose that will likely span the resulting landscape map. For each product you have expert descriptive mean data, a portion of which is shown in Table 1. In order to have the best likelihood of capturing all relevant drivers of liking and of spanning the future landscape map, you wish to find the twelve products among the eighteen that are most different from one another.

Independent Sets and Graph Theory: To simplify the problem of finding twelve maximally different products we first suppose that we have classified all pairs of products as either *similar* or *not-similar*. Our goal then becomes that of finding twelve products that are all *not-similar* to one another.

Product	Sweetness	Bitterness	Oiliness	
1	4.07	3.93	3.29	…
2	6.78	1.50	3.61	…
3	3.49	4.43	3.69	…
4	5.84	2.45	3.36	…
5	6.33	2.04	3.14	…
6	3.08	4.59	3.33	…
7	3.07	4.70	3.71	…
8	3.34	4.44	3.58	…
9	3.16	4.57	3.05	…
…	…	…	…	

Table 1. A portion of the expert descriptive data on the eighteen products (7-point scale).

In this simplified language, we see that when we represent our eighteen products visually, with products connected by a line segment when they are considered *similar*, we obtain a graph as shown in Figure 1. This perspective transforms our problem to one in graph theory, allowing us to use well developed mathematical tools[2,5].

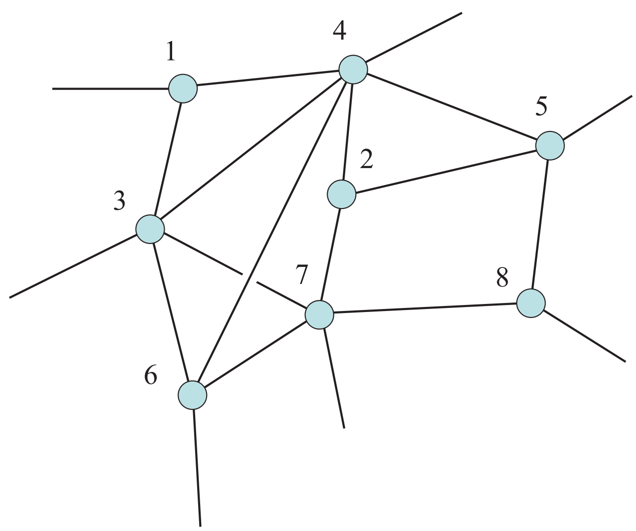

Figure 1. A graph representing similarity of product pairs.

Figure 1 shows eight of the eighteen products as nodes on a graph that are connected when the corresponding products are considered *similar*. Our goal now is to find twelve vertices that are all not connected to each other. This problem is well known in graph theory as the problem of finding independent sets. Although this problem is extremely computationally intensive in general, state-of-the-art algorithms typically allow us to find the complete collections of maximal independent sets within reasonable time for graphs with fewer than 100 nodes[3,4].

Similarity of Product Pairs: Returning to your problem, you now seek a way to classify product pairs as either *similar* or *not-similar*. For this you first compute the standardized distances between the products, obtaining a distance matrix, part of which is shown in Table 2. You then require a threshold distance to be applied to the distance matrix that yields at least one independent set of size twelve but none of size thirteen. You will use this threshold to declare products that are within the threshold of each other to be *similar* and those further apart than the threshold to be *not-similar*. Recent research on the theory of independent sets has facilitated solutions to problems of this type and they will be used here to find the best set of twelve products to choose for the category appraisal.

Product	1	2	3	4	…
1	0	7.5	2.2	4.5	…
2	7.5	0	8.1	5.2	…
3	2.2	8.1	0	4.9	…
4	4.5	5.2	4.9	0	…
…	…	…	…	…	…

Table 2. A portion of the standardized distance matrix for the eighteen products.

Determining an Optimal Collection of Products: Using this method you find a threshold for which there exists a collection of twelve products that are all *not-similar* to each other, but for which there is no such collection of thirteen products. Thus, you determine an optimal collection of twelve maximally different products for testing. This collection is shown in Figure 2 for four of the twelve products. The optimality of this solution avoids the arbitrariness introduced when other methods are used to address the same problem. In your case, there happened to be a unique collection, but more generally there could be multiple collections of maximally different products. In this case you could choose any of the sets of twelve, but most likely they would be chosen on the basis of other practical considerations such as availability, cost or competitiveness with the knowledge that you are selecting among equally acceptable optimal alternatives from the standpoint of the category landscape.

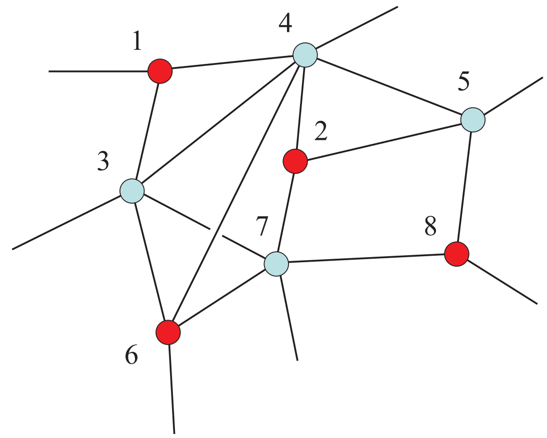

Figure 2. An independent set corresponding to the selected products (1,2,6,8).

You conduct your study with the twelve products determined above. Using the liking data for the twelve products, you conduct a Landscape Segmentation Analysis®. The twelve products span the space and the drivers of liking explain the space well. Using the expert descriptive information you place the remaining six products on the map and obtain predictions for these additional product locations as was shown in a previous technical report[1].

Conclusion: The choice of the products to be included in a product landscape study has a large effect on the future utility of the map and practical considerations limit the number of products that can reasonably be included in any such study. In many cases, techniques used to guide product selection have not consistently provided clear guidance. Using methods from the field of graph theory we now have a method that operates deterministically to provide clear guidance. This new method allows us to select a collection of products that are maximally different from each other and that will have the best chance of spanning the product landscape. Once the landscape is obtained, drivers of liking can be determined and the locations of other products not selected for original inclusion can be predicted. From this information, their acceptability to particular consumer segments can be estimated without incurring the additional cost of consumer testing.

References

1. Rousseau, B., and Ennis, D.M. (2008). Improving the cost and speed of product innovation. *IFPress* **11**(1), 2-3.
2. Valiente, G. (2002). *Algorithms on Trees and Graphs*. NY: Springer.
3. Lawler, E., Lenstra, J.K. and Kan, H.G.R. (1980). Generating all maximal independent sets: NP-hardness and polynomial-time algorithms. *SIAM Journal on Computing*. **9,** 558-565.
4. Cazals, F. and Karande, C. (2008). A note on the problem of reporting maximal cliques. *Theoretical Computer Science*. **407** (1-3), 564-568.
5. Ennis, J.M., Fayle, C.M. and Ennis, D.M., 2009. Reductions of letter displays. *IFPress Research Papers* **903,** 1-44.