Assortment Optimization
Under the General Luce Model

Alvaro Flores* Gerardo Berbeglia† Pascal Van Hentenryck‡

Tuesday 13th June, 2017

Abstract

This paper studies the assortment optimization problem under the General Luce Model (GLM), a discrete choice introduced by Echenique and Saito (2015) that generalizes the standard multinomial logit model (MNL). The GLM does not satisfy the Independence of Irrelevant Alternatives (IIA) property but it ensures that each product has an intrinsic utility and uses a dominance relation between products. Given a proposed assortment S, consumers first discard all dominated products in S before using a MNL model on the remaining products. The General Luce Model may violate the traditional regularity condition, which states that the probability of choosing a product cannot increase if the offer set is enlarged. As a result, the model can model behaviour that cannot be captured by any discrete choice model based on random utilities.

The paper proves that the assortment problem under the GLM is polynomially-solvable. Moreover, it proves that the capacitated assortment optimization problem under the General Luce Model (GLM) is NP-hard and presents polynomial-time algorithms for the cases where (1) the dominance relation is utility-correlated and (2) its transitive reduction is a forest. The proofs exploit a strong connection between assortments under the GLM and independent sets in comparability graphs.

1 Introduction

The optimal assortment problem is one of the most studied problems in revenue management: It consists in selecting a subset of products to offer customers in order to maximize revenue. Consider, for instance, an electronic store with limited space allocated to mobile phones. If the store has more than 500 mobile phones that can be acquired through its distributors (in various combinations of brands and sizes) and the mobile phone aisle has capacity to fit 50 cellphones on the shelves, the store manager has to decide which subset of products to offer given the product costs and the customer demand.

The optimal assortment problem requires a model predicting how customers select products when presented with a finite set of alternatives. Customer choice has been deeply studied in *discrete choice theory*, which associates a *utility* to each alternative. Different assumptions on the distribution of the utilities lead to different discrete choice models: Celebrated examples include the

*College of Engineering & Computer Science, Australian National University.

†Melbourne Business School, The University of Melbourne.

‡Industrial and Operations Engineering & Computer Science and Engineering, University of Michigan, Ann Arbor (pvanhent@umich.edu).

multinomial logit (MNL) (Luce, 1959), the mixed multinomial logit (MMNL) (Daly and Zachary, 1978), and the nested multinomial logit (NMNL) (Williams, 1977).

The multinomial logit model (MNL), also known as the Luce model, is widely used in discrete choice theory. Since the model was introduced by Luce (1959), it was applied to a wide variety of demand estimation problems arising in transportation (McFadden, 1978; Catalano, Lo Casto, and Migliore, 2008), marketing (Guadagni and Little, 1983; Gensch, 1985; Rusmevichientong, Shen, and Shmoys, 2010), and revenue management (Talluri and Van Ryzin, 2004; Rusmevichientong, Shen, and Shmoys, 2010). One of the reasons for its success stems from its small number of parameters (one for each product): This allows for simple estimation procedures with no overfit, even when there is limited historical data (McFadden, 1974). Howver, one of the flaws of the multinomial logit model is the property known as the *Independence of Irrelevant Alternatives (IIA)*, which states that the ratio between the probabilities of choosing elements x and y is constant regardless of the offered subset. This property does not hold when products cannibalize each other or are perfect substitutes (Ben-Akiva and Lerman, 1985; Debreu, 1960; Anderson, Depalma, and Thisse, 1992).

Several extensions to the MNL model have been introduced to overcome the IIA property and some of its other weaknesses; They include the nested multinomial logit and the latent class MNL model. These models however do not handle zero-probability choices well. Consider two products a and b: The MNL model states that the probability of selecting a over b depends on the relative utility of a compared to the utility of b. Consider the case in which b is never selected when a is offered. Under the MNL model, this means that b must have zero utility. But this would prevent b from being selected even when a is not offered in an assortment.

This paper considers the General Luce Model (GLM) that was recently introduced by Echenique and Saito (2015) and which allows for violations to the IIA property. The General Luce Model generalizes the MNL by incorporating a dominance (anti-symmetric and transitive) relation among the alternatives. Under such relationship, the presence of an alternative x may prevent another alternative y from being chosen despite the fact that both are present in the offered assortment. In this case, alternative x is said to *dominate* alternative y. However, when x is not present, y might be chosen with positive probability if it is not dominated by any other product z.

An important application of the GLM can be found in assortment problems where there exists a direct way to compare the products over a set of features. For illustration, consider a telecommunication company offering phone plans to consumers. A plan is characterized by a set of features such as price per month, free minutes in peak hours, free minutes in weekends, free data, price for additional data, and price per minute to foreign countries. Given two plans x and y, we say that plan x *dominates* plan y, if the price per month of x is less than that of y, and x is *at least* as good as y in every single feature. In the past, the company offered consumers a certain set of plans S_t each month t such that no plan in S_t is dominated by another plan (in S_t). The offered plans however were different each month. Using historical data and assuming that consumers preferences can be approximated using a multinomial logit, it is possible to perform a robust estimation procedure to obtain the parameters of such MNL model. Once the parameters are obtained, the assortment problem consists in finding the best assortment of phones plans S^* to maximize the expected revenue. A natural constraint in this problem consisting in enforcing that every phone plan offered in S^* cannot be dominated by any other. Proposition 1 in Section 3 shows that the problem discussed here can be reduced to solving assortment problem under the GLM.

The contributions of the paper can be summarized as follows. The first key contribution of this paper is to show that the assortment problem under the GLM can be solved in polynomial time. This result is particularly interesting since the GLM may violate the traditional regularity condition, which states that the probability of choosing a product cannot increase if the offer set is enlarged. The proof uses two main results: the polynomial-time solvability of the maximum-independent set in a comparability graph (Möhring, 1985) and a seminal result by Megiddo (1979). The second key contribution is to show that the capacitated assortment problem under the GLM is NP-hard, which contrasts with results on the MNL. Finally, the the proposed polynomial algorithms for two interesting subcases of the capacitated assortment problem: (1) When the dominance relation is utility-correlated and (2) when the transitive reduction of the dominance relation can be represented as a forest. The proofs use a strong connection between assortments under the GLM and independent sets.

The rest of the paper is organized as follows: Section 2 presents a review of the literature. Section 3 formalizes the GLM and some of its properties. Section 4 proves that assortment optimization under the GLM is polynomial-time solvable. Section 5 presents the results on the capacitated version. Section 6 concludes the paper.

2 Literature Review

Since the assortment problem is a very active research topic, we focus on recent results closely related with this paper and, in particular, assortment problems over the multinomial logit model (MNL) (Luce, 1959) and the development of new discrete choice models.

Despite the IIA property, the MNL is widely used. Indeed, for many applications, the mean utility of a product can be modeled as a linear combination of its features. If the features capture the mean utility associated with each product, then the error between the utilities and their means may be considered as independent noise and the MNL emerges as a natural candidate for modeling customer choice. In addition, the MNL parameters can be estimated from customer choice data, even with limited data (Ford, 1957; Negahban, Oh, and Shah, 2012), because the associated estimation problem has a concave log likelihood function (McFadden, 1974) and it is possible to measure how good the fitted MNL approximates the data (Hausman and McFadden, 1984). Moreover, it is possible to improve model estimation when the IIA property is likely to be satisfied (Train, 2003). Recently, Jagabathula and Vulcano (2015) proposed a partial-order model to estimate individual preferences, where preference over products are modeled using forests. They cluster the customers in classes, each class being represented with a forest. When facing an assortment S, customers select, following an MNL model, products that are roots of the forest projected on S. This approach outperformed state-of-the-art methods when measuring the accuracy of individual predictions.

One of the first positive results on the assortment problem under the multinomial logit model was obtained by Talluri and Van Ryzin (2004), where the authors showed that the optimal assortment can be found by greedily by adding products to the offered assortment in the order of decreasing revenues, thus evaluating at most a linear number of subsets. Rusmevichientong, Shen, and Shmoys (2010) studied the assortment problem under the MNL but with a capacity constraint limiting the products that can be offered. Under these conditions, the optimal solution is not necessarily a revenue-ordered assortment but it can still be found in polynomial time.

Gallego, Ratliff, and Shebalov (2011) proposed a more general attraction model where the probabilities of choosing a product depend on all the products (not only the offered subset as in the MNL). This involves a shadow attraction value associated with each product that influence the choice probabilities when the product is not offered. Davis, Gallego, and Topaloglu (2013) showed that a slight transformation of the MNL model allows for the solving of the assortment problem when the choice probabilities follow this more sophisticated attraction model. This continues to hold when assortments must satisfy a set of totally unimodular constraints.

The Mixed Multinomial Logit (Daly and Zachary, 1978) is an extension of the MNL model, where different sets of customers follow different MNL models. Under this setting, the problem becomes NP-hard (Bront, Méndez-Díaz, and Vulcano, 2009) and it remains NP-hard even for two customer types (Rusmevichientong et al., 2014). A branch-and-cut algorithm was proposed by Méndez-Díaz et al. (2014). Feldman and Topaloglu (2015) proposed methods to obtain good upper bounds on the optimal revenue. Rusmevichientong and Topaloglu (2012) considered a model where customers follow a MNL model and the parameters belong to a compact uncertainty set. The firm wants to hedge against the worst-case scenario and the problem amounts to finding an optimal assortment under this uncertainty conditions. Surprisingly, when there is no capacity constraint, the revenue-ordered strategy is optimal in this setting. Jagabathula (2014) proposed a local-search heuristic for the assortment problem under an arbitrary discrete choice model. Davis, Gallego, and Topaloglu (2013) and Abeliuk et al. (2016) proposed polynomial time algorithms to solve the assortment problem under the MNL model with capacity constraint and position bias, where position bias means that customer choices are affected by the positioning of the products in the assortment.

Attention has also been devoted to discrete choice models to represent customer choices in more realistic ways, including models that violate the IIA property (Ben-Akiva and Lerman, 1985). This property does not always hold in practice (Rieskamp, Busemeyer, and Mellers, 2006), including when products cannibalize each other (Ben-Akiva and Lerman, 1985). Echenique, Saito, and Tserenjigmid (2013) identify these violations as perception priorities, and adjust probabilities to take their effects into account. Gul, Natenzon, and Pesendorfer (2014) provide an axiomatic generalization of MNL model to address the case where the products share features. Fudenberg and Strzalecki (2015) propose an axiomatic generalization of a discounted logit model incorporating a parameter to model the influence of the assortment size.

Several models have been proposed to address the issue of zero-probability choices. Masatlioglu, Nakajima, and Ozbay (2012) propose a theoretical foundation for maximizing a single preference under limited attention, i.e., when customers select among the alternatives that they pay attention to. Manzini and Mariotti (2014) incorporate the role of attention into stochastic choice, proposing a model in which customers consider each offered alternative with a probability and choose the alternative maximizing a preference relation within the considered alternatives. This was axiomatized and generalized in Brady and Rehbeck (2016), by introducing the concept of *random conditional choice set rule*, which captures correlations in the availability of alternatives. This concept also provided a natural way to model substitutability and complementarity.

The General Luce Model considered in this paper was proposed by Echenique and Saito (2015); It handles zero-probability choice by introducing the concept of *dominance*, meaning that if a product x dominates a product y, then y is never selected in presence of y. Our main purpose is to study the computational issues arising in finding optimal assortments under this model.

3 The General Luce Model

The GLM overcomes a key limitation of the MNL: The fact that a product must have zero utility if it has zero probability to be chosen in a particular assortment. This limitation means that the product cannot be chosen with positive probability in any other assortment. The GLM eliminates this pathological case through the concept of *consideration function* which, given a set of products S, returns a subset of S with positive probability of being selected. Let X denotes the set of all products and let $u(x) \geq 0$ be the utility of product $x \in X$. Given an assortment $A \subseteq X$, a stochastic choice function ρ returns a probability distribution over A, i.e., $\rho(x, A)$ is the probability of picking x in the assortment A.

Definition 1 (General Luce Model)**.** The choice function ρ is a General Luce Model if there exists a utility function $u : X \to \mathbb{R}^+$ and a function $c : 2^X \setminus \emptyset \to 2^X \setminus \emptyset$ with $c(A) \subseteq A$ for all $A \subseteq X$ such that

$$\rho(x, A) = \begin{cases} \frac{u(x)}{\sum_{y \in c(A)} u(y)} & \text{if } x \in c(A), \\ 0 & \text{if } x \notin c(A). \end{cases} \tag{1}$$

and the function c satisfies the following properties:

1. $A \subseteq B \implies c(B) \cap A \subseteq c(A)$.
2. $\forall A' \subset A$ with $x \in A' : x \in c(A') \implies x \in c(A)$.
3. $x \notin c(A) \implies c(A) = c(A \setminus \{x\})$.

Property 1 ensures that, if a product is selected with positive probability from a set B where it faces more competition than in $A \subseteq B$, then the product also has positive probability of being chosen in A. Property 2 states that, if product x has positive probability of being chosen over every subset of A containing x, then x has also positive probability of being chosen over A. Property 3 states that, if x cannot be chosen with positive probability in an assortment A, then it cannot affect the selection of other objects in A. It is useful to mention two interesting special cases:

1. If $c(S)$ is a singleton for all $S \subseteq X$, then $\rho(x, S)$ is a deterministic choice.
2. If $c(S) = S$ for all $S \subseteq X$, then the GLM coincides with the MNL.

To solve the assortment problem under the GLM, we use an important result by Echenique and Saito (2015) that characterizes function c using an antisymmetric and transitive relation over X.

Definition 2. A function $c : 2^X \setminus \emptyset \to 2^X \setminus \emptyset$ satisfying $c(A) \subseteq A$ for all $A \subset X$ is **dominance rationalizable** if there is an antisymmetric and transitive relation $\succ$ such that:

$$c(A) = \{x \in A \mid \nexists y \in A : y \succ x\}. \tag{2}$$

Proposition 1. *[Echenique and Saito (2015)] A pair (u, c) is a GLM if and only if c is dominance rationalizable.*

Without loss of generality, we assume that the relation $\succ$ is irreflexive: if $x \succ x$, then product x can never be selected and is thus irrelevant. As a result, we can describe any GLM by an irreflexive, transitive, and antisymmetric relation $\succ$ that fully captures the relation between products. We also extend the utility function to consider the outside option, with index 0 and $u(0) \geq 0$, to model the fact that customers may not select any product. As result, the utility function has signature $u : X \cup \{0\} \to \mathbb{R}^+$ and the probability of purchasing a product $x \in S$ when offered the set S is given by:

$$\rho(x, S) = \frac{u(x)}{\sum_{y \in c(S)} u(y) + u_0}. \tag{3}$$

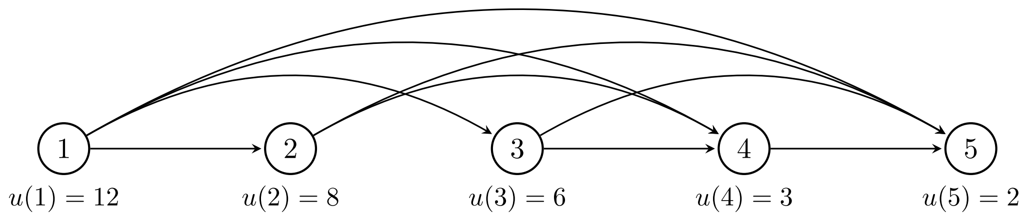

Figure 1: Example of a DAG for a General Threshold Luce Model

If $x \notin S$, then $\rho(x, S) = 0$. Note also that $\forall x \in S, \rho(x, S) = \rho(x, c(S))$.

Echenique and Saito (2015) also proposed a model called the *Threshold Luce Model*, where they explain dominance in terms of how big the utilities are when compared with each other. For a given threshold $t > 0$, they define the consideration set $c(S)$ for a set $S \subseteq X$ as:

$$c(S) = \{y \in S \mid \nexists x \in S : u(x) > (1 + t)u(y)\}. \tag{4}$$

Intuitively, a utility ratio of more than $(1+t)$ means that the less-preferred alternative is dominated by the more-preferred alternative. In other words, $x \succ y$ if and only if $u(x) > (1 + t)u(y)$. Observe that the relation $\succ$ is clearly irreflexive, transitive, and antisymmetric.

Example 1. [**Feature Difference Threshold:**] Assume that each product has a set of features $\mathcal{F} = \{1, \ldots, m\}$. A product x can then be represented by a m-dimensional vector $x \in \mathbb{R}^m$. Assume that the perceived relevance of each feature k is measured by a weight ν_k. The dominance relation can be defined as $x \succ y \iff \sum_{k=1}^{m} \nu_k(x_k - y_k) \geq T$, where $T > 0$ is a tolerance parameter that represents how much difference a customer allows before considering that an alternative dominates another. The dominance relation is irreflexive, transitive, and antisymmetric and hence it can be used to define an instance of the GLM.

The dominance relation $\succ$ can be represented as a Directed Acyclic Graph (DAG), where the nodes represent the products and there is a directed edge (x, y) if and only if $x \succ y$. Sets satisfying $c(S) = S$ are *anti-chains* in the DAG, meaning that there are no arc connecting them.

Example 2 (DAG Representation of a GLM)**.** Consider the *Threshold Luce Model* defined over $X = \{1, 2, 3, 4, 5\}$ with utilities $u(1) = 12, u(2) = 8, u(3) = 6, u(4) = 3$ and $u(5) = 2$, and threshold $t = 0.4$. We have that $i \succ j$ iff $u(j) > 1.4\, u(j)$ and the DAG is depicted in Figure 1.

4 Assortment Under the General Luce Model

This section studies the assortment problem for the GLM using the definitions and notations presented earlier. Let $r : X \cup \{0\} \to \mathbb{R}^+$ be a revenue function associated with each product and satisfying $r(0) = 0$. The expected revenue of a set $S \subseteq X$ is given by

$$R(S) = \sum_{i \in c(S)} \rho(i, S) r(i). \tag{5}$$

The assortment problem amounts to finding a set

$$S^* \in \operatorname*{argmax}_{S \subseteq X} R(S)$$

giving an optimal assortment revenue

$$R^* = \max_{S \subseteq X} R(S).$$

Every subset $S \subseteq X$ can be uniquely represented by a binary vector $x \in \{0,1\}^n$ such that $i \in S$ if and only if $x_i = 1$. Using this bijection, the search space for S^* can be restricted to

$$\mathcal{D} = \{x \in \{0,1\}^n \mid \; \forall s \succ t : x_s + x_t \leq 1\}$$

where $\mathcal{D}$ represents all the subsets satisfying $S = c(S)$, which means that no product on S dominates another product in S. There is always an optimal solution S^* that belongs to $\mathcal{D}$ because $R(S) = R(c(S))$ and $c(S) \in D$ for all sets S in X. As a result, the Assortment Problem under the GLM (AP-GLM) can be formulated as

$$\begin{aligned} \underset{x}{\text{maximize}} \quad & \frac{\sum_{i=1}^n r_i u_i x_i}{\sum_{i=1}^n u_i x_i + u_0} \\ \text{subject to} \quad & x \in \mathcal{D} \end{aligned} \qquad \texttt{(AP-GLM)}$$

where r_i and u_i represent $r(i)$ and $u(i)$ for simplicity.

An effective strategy for solving many assortment problems consists in considering revenue-ordered assortments, which are obtained by choosing a threshold ρ and selecting all the products with revenue at least ρ. This strategy leads to an optimal algorithm for the assortment problem under the MNL. Unfortunately, it fails under the GLM because adding a high-utility product may remove many dominated products whose utilities would lead to a higher revenue.

Example 3 (Sub-Optimality of Revenue-Ordered Assortments)**.** Consider a Threshold Luce Model with $X = \{1,2,3\}$, revenues $r(1) = 88, r(2) = 47, r(3) = 46$, utilities $u(0) = 55, u(1) = 13, u(2) = 26, u(3) = 15$ and $t = 0.6$. Then $x \succ y$ iff $u(x) > 1.6\ u(y)$ which gives $2 \succ 1$ and $2 \succ 3$. Consider the sets $S \subseteq X$ satisfying $S = c(S)$:

S	$R(S)$
$\{1\}$	16.824
$\{2\}$	15.086
$\{3\}$	9.857
$\{1,3\}$	22.096

The optimal revenue is given by assortment $\{1,3\}$, while best revenue-ordered assortment under the GLM is $S = \{1\}$.

Our next example indicates that the GLM may also violate regularity conditions.

Example 4. [Violation of The Regularity Condition] This example shows how the GLM may violate the regularity condition. Consider $X = \{1,2,3\}$, the utilities are $u(1) = 1, u(2) = 1, u(3) = 2$, the utility of the outside option is $u(0) = 1$ and the only dominance relation is $2 \succ 3$. Whenever product 2 is offered, product 3 is never selected. Consider the following assortments $S = \{1,3\}$ and $S' = \{1,2,3\}$. According the GLM, the probability of selecting product 1 on each of these assortments is:

$$\rho(1,S) = \frac{u(1)}{u(1) + u(3) + u(0)} = \frac{1}{4} = 0.25$$

$$\rho(1, S') = \frac{u(1)}{u(1) + u(2) + u(0)} = \frac{1}{3} = 0.\overline{3}$$

We have that $S \subset S'$ but $\rho(1, S) < \rho(1, S')$, which violates the regularity condition.

To solve problem `AP-GLM`, consider first the `MaxUtility` problem defined over the same set of constraints. Given weights $c_i \in \mathbb{R}$ $(1 \leq i \leq n)$, the `MaxUtility` problem is defined as follows:

$$\begin{aligned} \underset{x}{\text{maximize}} \quad & \sum_{i=1}^{n} c_i x_i \\ \text{subject to} \quad & x \in \mathcal{D} \end{aligned} \tag{MaxUtility}$$

We now show that (`MaxUtility`) can be reduced to the maximum weighted independent set problem in a directed acyclic graph with positive vertex weights. An independent set is a set of vertices I such there is no edge connecting any two vertices in I. The maximum weighted independent set problem (`MWIS`) can be stated as follows:

Definition 3. *Maximum Weighted Independent Set Problem:* Given a graph $G = (V, E)$ with a weight function $w : V \to \mathbb{R}$, find an independent set $I^* \in \text{argmax}_{I \in \mathcal{I}} \sum_{i \in I} w(i)$, where $\mathcal{I}$ is the set of all independent sets.

Recall that the dominance relation can be represented as a DAG G which includes an arc (u, v) whenever $u \succ v$. As a result, the condition $x \in \mathcal{D}$ implies that any feasible solution to (`MaxUtility`) represents an independent set in G and maximizing $\sum_{i=1}^{n} c_i x_i$ amounts to finding the independent set maximizing the sum of the weights. Since the dominance relation is a partial order, the DAG representing the dominance relation is a comparability graph and the following result is particularly interesting.

Theorem 1 (Möhring (1985))**.** *The maximum weighted independent set is polynomially-solvable for comparability graphs with positive weights.*

We are ready to present our first result.

Lemma 1. (`MaxUtility`) *is polynomial-time solvable.*

Proof. We first show that we can ignore those products with a negative weight. Let $\hat{X} = \{i \in X \mid c_i > 0\}$ and $\hat{\mathcal{D}} = \{x \in \{0,1\}^n \mid \forall s, t \in \hat{X}, s \succ t : \; x_s + x_t \leq 1\}$. Solving (`MaxUtility`) is equivalent to solving:

$$\begin{aligned} \underset{x}{\text{maximize}} \quad & \sum_{i \in \hat{X}} c_i x_i \\ \text{subject to} \quad & x \in \hat{\mathcal{D}} \end{aligned} \tag{Reduced MaxUtility}$$

Indeed, consider an optimal solution x^* to Problem `MaxUtility` and assume that there exists $i \in X$ such that $c_i < 0$ and $x_i^* = 1$. Define $\hat{x}$ like x^* but with $\hat{x}_i = 0$. $\hat{x}$ has a strictly better revenue than x^* and is feasible since setting a component to zero cannot violate any constraint (i.e., $\hat{x} \in \mathcal{D}$). This contradicts the optimality of x^*. Now Problem `Reduced MaxUtility` can be reduced to solving an instance of Problem `MWIS` in a DAG with positive weights that corresponds to the dominance relation. This DAG is a comparability graph and the result follows from Theorem 1. $\square$

The next step in solving the assortment problem under the GLM relies on a result by Megiddo Megiddo (1979). Let D be a domain defined by some set of constraints and consider Problem $\mathtt{A}$

$$\begin{aligned} \underset{x}{\text{maximize}} \quad & \sum_{i=1}^{n} c_i x_i \\ \text{subject to} \quad & x \in D \end{aligned} \tag{A}$$

and its associated Problem $\mathtt{B}$:

$$\begin{aligned} \underset{x}{\text{maximize}} \quad & \frac{a_0 + \sum_{i=1}^{n} a_i x_i}{b_0 + \sum_{i=1}^{n} b_i x_i} \\ \text{subject to} \quad & x \in D. \end{aligned} \tag{B}$$

Using this notation, Megiddo's theorem can be stated as follows.

Theorem 2 (Megiddo (1979))**.** *If Problem* $\mathtt{A}$ *is solvable within* $O(p(n))$ *comparisons and* $O(q(n))$ *additions, then Problem* $\mathtt{B}$ *is solvable in* $O(p(n)(q(n) + p(n)))$ *time.*

We are now in position to state our main theorem.

Theorem 3. *The assortment problem under the General Luce Model is polynomial-time solvable.*

Proof. Recall that the assortment problem under the GLM (`AP-GLM`) can be formulated as

$$\begin{aligned} \underset{x}{\text{maximize}} \quad & \frac{\sum_{i=1}^{n} r_i u_i x_i}{\sum_{i=1}^{n} u_i x_i + u_0} \\ \text{subject to} \quad & x \in \mathcal{D} \end{aligned} \tag{6}$$

where $\mathcal{D} = \{x \in \{0,1\}^n \mid \forall s \succ t : x_s + x_t \leq 1\}$.

The problem of maximizing the numerator in (6) is exactly the `MaxUtility` problem. By Lemma 1, this is polynomial-time solvable. Now observe that (6) (i.e., problem `AP-GLM`) can be seen as a Problem $\mathtt{B}$. Therefore, by Theorem 2, the assortment problem under the GLM is solvable in polynomial time. $\square$

In addition to solving the assortment problem under the GLM, Theorem 3 is interesting in that it solves the assortment problem under a Multinomial Logit with a specific class of constraints. It can be contrasted with the results by Davis, Gallego, and Topaloglu (2013), where feasible assortments satisfy a set of totally unimodular constraints. They show that the resulting problem can be solved as a linear program. However, the GLM introduces constraints that are not necessarily totally unimodular as we now show.

Example 5. Consider $X = \{1, 2, 3, 4\}$ and $1 \succ 3, 1 \succ 4, 2 \succ 3, 2 \succ 4$, and $3 \succ 4$. The constraint matrix that defines the feasible space ($\mathcal{D}$) for this instance is:

$$M = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

where each row represents a constraint $x_u + x_v \leq 1$. meaning that just one end of the edge can be selected at the time. Camion (1965) proved that M is totally unimodular if and only if, for every (square) Eulerian submatrix A of M, $\sum_{i,j} a_{ij} \equiv 0 \pmod 4$. Consider the sub-matrix corresponding to the first, second, and fifth rows and the first, third, and fourth columns

$$N = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

Matrix N is eulerian (The sums of every element on each row or on each column is a multiple of 2). But the sum of all elements of N is $6 \not\equiv 0 \pmod 4$ and hence M is not totally unimodular.

5 The Capacitated Assortment Problem

In many applications, the number of products in an assortment is limited, giving rise to capacitated assortment problems. Let C $(1 \leq C \leq n)$ be the maximum number of products allowed in an assortment. The Capacitated Assortment Problem under the General Luce Model (`CGLMAP`) is given by

$$\begin{aligned} \underset{x}{\text{maximize}} \quad & \frac{\sum_{i=1}^{n} r_i u_i x_i}{\sum_{i=1}^{n} u_i x_i + u_0} \\ \text{subject to} \quad & x \in \mathcal{D}_C \end{aligned} \tag{CGLMAP}$$

where $\mathcal{D}_C = \{x \in \{0,1\}^n \mid \forall (s,t) \in \mathcal{R} \quad x_s + x_t \leq 1 \wedge \sum_{i=1}^{n} x_i \leq C\}$. As before, it is useful to define its maximum-utility counterpart, i.e.,

$$\begin{aligned} \underset{x}{\text{maximize}} \quad & \sum_{i=1}^{n} c_i x_i \\ \text{subject to} \quad & x \in \mathcal{D}_C \end{aligned} \tag{MUCC}$$

This section first proves that the capacitated assortment problem under the GLM is NP-hard. The reduction uses the Maximum Weighted Budgeted Independent Set (`MWBIS`) problem proposed by Bandyapadhyay (2014) which amounts to finding a maximum weighted independent set of size not greater than C. Kalra et al. (2017) showed that Problem (`MWBIS`) is NP-hard for bipartite graphs.

Theorem 4. *Problem* (`CGLMAP`) *is NP-hard (under Turing reductions).*

Proof. The proof considers four problems:

1. Problem (`MWBISBP`): Maximum weighted independent set of size at most than C for bipartite graphs.

2. Problem (`MWEBISBP`): Maximum weighted independent set of size equal to C for bipartite graphs.

3. Problem (`ECGLMAP`): Optimal assortment under the General Luce Model of size C.

4. Problem (`CGLMAP`): Optimal capacitated assortment under the General Luce Model of size at most C.

The proof shows that Problems (`MWEBISBP`), (`ECGLMAP`), and (`CGLMAP`) are NP-hard, using the NP-hardness of Problem (`MWBISBP`) (Kalra et al., 2017) as a starting point.

First observe that Problem (`MWEBISBP`) is NP-hard under Turing reductions. Indeed, Problem (`MWBISBP`) can be reduced to solving C instances of Problem (`MWEBISBP`) with budget c ($1 \leq c \leq C$).

We now show that Problem (`ECGLMAP`) is NP-hard. Consider Problem (`MWEBISBP`) over a bipartite graph $G = (V = V_1 \cup V_2, E)$, where $V_1 \cap V_2 = \emptyset$, every edge $(v_1, v_2) \in E$ satisfies $v_1 \in V_1$ and $v_2 \in V_2$, w_v is the weight of vertex v, and C is the budget. We show that Problem (`MWEBISBP`) over this bipartite graph can be polynomially reduced to Problem (`ECGLMAP`). The reduction assigns each vertex v to a product with $u(v) = 1$ and $r_v = w_v$, sets $u_0 = 0$, and has a capacity C. Moreover, the reduction uses the following dominance relation: $v_1 \succ v_2$ iff $(v_1, v_2) \in E$. This dominance relation is irreflexive, anti-symmetric, and transitive, since the graph is bipartite. A solution to Problem (`MWEBISBP`) is a feasible solution to Problem (`ECGLMAP`), since the independent set cannot contain two vertices v_1, v_2 with $v_1 \succ v_2$ by construction. Similarly, a feasible assortment is an independent set, since the assortment cannot select two vertices $v_1 \in V_1$ and $v_2 \in V_2$ with $(v_1, v_2) \in E$, since $v_1 \succ v_2$. The objective function of Problem (`ECGLMAP`) reduces to maximizing

$$\frac{1}{C} \sum_{v \in V} r_v x_v$$

which is equivalent to maximizing $\sum_{v \in V} r_v x_v$ since exactly C products will be selected by every feasible assortment. The result follows by the NP-hardness of Problem (`MWEBISBP`).

Finally, Problem (`CGLMAP`) is NP-hard under Turing reductions. Indeed, Problem (`CGLMAP`) can be reduced to solving C instances of Problem (`ECGLMAP`) with capacity c ($1 \leq c \leq C$). $\square$

It is interesting to mention that Problem (`MUCC`) is equivalent to finding an anti-chain of maximum weight among those of cardinality at most C. This problem (`MWLA`) was proposed by Shum and Trotter (1996) and its complexity was left open, but the above results show that it is also NP-hard. Bandyapadhyay (2014) studied Problem (`MWBIS`) for various types of graphs (e.g., trees and forests), but the dominance relation of the GLM can never be a tree since it is transitive.

In light of this NP-hardness result, the rest of this section presents polynomial-time algorithms for two special cases of the dominance relation.

5.1 The Utility-Correlated General Luce Model

The first special case considers a dominance relation that is correlated with utilities.

Definition 4 (Utility-Correlated General Luce Model)**.** A General Luce Model is utility-correlated if the dominance relation satisfies the following two conditions:

1. If $x \succ y$, then $u(x) > u(y)$.
2. If $x \succ y$ and $u(z) > u(x)$, then $z \succ y$.

The first condition simply expresses that product x can only dominate product y if the utility of x is greater than the utility of y. The second condition ensures that, if x dominates y, then any product whose utility is greater than x also dominates y. The induced dominance relation is irreflexive, anti-symmetric, and transitive.

The capacitated assortment optimization can be solved in polynomial time under the Utility-Correlated General Luce Model. Consider an assortment whose product with the largest utility is

Algorithm 1: Capacitated Assortment Optimization under the Utility-Correlated GLM.

Data: $X, \succ, r, u$
Result: Optimal Assortment S^*
$R(S^*) = 0$ **for** $k = 1, \ldots, n$ **do**
  $X_k = \{i \in X \mid u(i) \leq u(k) \ \& \ k \not\succ i\}$
  $S_k =$ `CMLMAP`(X_k, r, u)
  if $R(S_k) > R(S^*)$ **then**
    $S^* = S_k$
  end
end
return S^*

k. This assortment cannot contain any product dominated by k. Moreover, if k_1 and k_2 are two other products in this assortment, then k_1 cannot dominate k_2 since k would also dominate k_2. As a result, consider the set

$$X_k = \{i \in X \mid u(i) \leq u(k) \ \& \ k \not\succ i\}.$$

No product in X_k dominates any other product in X_k and hence the `CGLMAP` reduces to a traditional assortment problem under the MNL. This idea is formalized in Algorithm 1, where `CMLMAP` is a traditional algorithm for the MNL. The algorithm considers each product in turn and the products that it does not dominate and applies a traditional capacitated assortment optimization under the MNL. The best such assortment is the solution to the capacitated assortment under the utility-correlated GLM.

Theorem 5. *`CGLMAP` can be solved in polynomial time under the Utility-Correlated GLM.*

Proof. To show correctness, it suffices to show that the optimal assortment must be a subset of one of the X_k $(1 \leq k \leq n)$. Let A be the optimal assortment and assume that k is its product with the largest utility (break ties randomly). A must be included in X_k since otherwise it would contain a product x such that $k \succ x$ (contradicting feasibility) or such that $u(x) > u(k)$ (contradicting our hypothesis). The correctness then follows since there is no dominance relationship between any two elements in each of X_k. The claim of polynomial-time solvability follows from the availability of polynomial-time algorithms for the assortment problem under the MNL and the fact that are exactly n calls to such an algorithm. □

Note also that the result can also be generalized to capacitated assortment problem with position optimization (e.g., (Davis, Gallego, and Topaloglu, 2013; Abeliuk et al., 2016)).

5.2 The General Luce Model over Tree-Induced Dominance Relations

Let $\mathcal{R}_\succ$ be the transitive reduction of the irreflexive, antisymmetric, and transitive relation $\succ$. This section considers the capacitated assortment problem when the relation $\mathcal{R}_\succ$ can be represented as a tree. Without loss of generality, we can assume that the tree contains all products. Otherwise, we can add another product with zero weight that dominates all original products. This new product will be the root of the tree and the products not in the original tree will be the children of the root. Similarly, the same transformation applies to the case when $\mathcal{R}_\succ$ is a forest. Here all the trees in the forest will be children of the new product.

We show how to solve Problem (`MUCC`). The result follows again by applying Megiddo's theorem. The first step of the algorithm simply removes all products with negative weight: Their children can be added to the parent of the deleted vertex. The main step then solves (`MUCC`) bottom-up using dynamic programming from the leaves. For simplicity, we present the recurrence relations to compute the weight of the optimal assortment. It is easy to recover the optimal assortment itself. The recurrence relations compute two functions:

1. $\mathcal{A}(k, c)$ which returns the weight of an optimal assortment using product k and its descendants in the tree representation of $\mathcal{R}_{\succ}$ for a capacity c;

2. $\mathcal{A}^{+}(S, c)$ which, given a set S of vertices that are children of a vertex k, returns the weight of an optimal assortment using the products in S and their descendants for a capacity c.

The key intuition behind the recurrence is as follows. If v is a vertex and v_1 and v_2 are two of its children, v_1 does not dominate v_2 or any of its descendants. Hence, it suffices to compute the best assortments producing $\mathcal{A}(v_1, 0), \ldots, \mathcal{A}(v_1, C)$ and $\mathcal{A}(v_2, 0), \ldots, \mathcal{A}(v_2, C)$ and to combine them optimally. The recurrence relations are defined as follows ($v \in X$ and $1 \le c \le C$):

$$\begin{aligned} &\mathcal{A}(v, 0) = 0; \\ &\mathcal{A}(v, c) = \max(c_v, \mathcal{A}^{+}(\mathit{children}(v), c)); \end{aligned}$$

and

$$\begin{aligned} &\mathcal{A}^{+}(\emptyset, 0) = 0; \\ &\mathcal{A}^{+}(S, c) = \max_{\substack{n_1, n_2 \ge 0 \\ n_1 + n_2 = c}} \mathcal{A}^{+}(S \setminus \{e\}, n_1) + \mathcal{A}(e, n_2) \text{ where } e = \operatorname*{argmax}_{i \in S} c_i. \end{aligned}$$

where *children(p)* denotes the children of product p in the tree. Note that $\mathcal{A}^{+}(S, c)$ is computed recursively to obtain the best assortment from the products in S and their descendants.

Theorem 6. *Let $\succ$ a dominance relation whose relation $\mathcal{R}_{\succ}$ is a tree containing all products. The capacitated assortment problem under the GLM and $\succ$ is polynomial-time solvable.*

Proof. By Theorem 2, it suffices to show that Problem (`MUCC`) is solved by the recurrences in polynomial time. The correctness of recurrence $\mathcal{A}(v, c)$ comes from the fact that vertex v dominates all its descendants and cannot be present in any assortment featuring any of them. The correctness of recurrence $\mathcal{A}^{+}(S, c)$ follows from the fact that e is not dominated by, and does not dominate, any element in S, since they are all children of the same node. This also holds for the descendants of e and the descendants of the elements in S. Hence, the optimal assortment is obtained by splitting the capacity c into n_1 and n_2 and merging the best assortment for $\mathcal{A}^{+}(S, n_1)$ and $\mathcal{A}(e, n_2)$ for some $n_1, n_2 \ge 0$ summing to c. The recurrences can be solved in polynomial time since the computation for each vertex v and capacity c takes $O(n\ C)$ time, giving an overall time complexity of $O(n^2\ C^2)$. □

6 Conclusion and Future Work

This paper studies the assortment optimization problem under the General Luce Model (GLM), a discrete choice model introduced by Echenique and Saito (2015) that generalizes the standard multinomial logit model (MNL) with a dominance relation and does not satisfy the Independence of Irrelevant Alternatives (IIA) property. The paper proved that the assortment problem under the GLM can be solved in polynomial time. The paper also considered the capacitated assortment problem under the GLM and proved that the problem becomes NP-hard in this setting. Finally, the paper presented polynomial-time algorithms for special cases of the capacitated problem when (1) the dominance relation is utility-correlated and when (2) its transitive reduction is a forest.

There are at least two interesting avenues for future research. First, one may wish to study how to generalize the GLM further while still keeping the assortment problem solvable in polynomial time. For example, one can try to check whether there exists a model that unifies the GML and the elegant work in Davis, Gallego, and Topaloglu (2013) where the assortment problem is still solvable in polynomial time. Second, given that capacitated version of the GLM is NP-hard under Turing reductions (Theorem 4), it is interesting to see whether there exist good approximation algorithms for this problem.

7 Acknowledgments

We thanks Yuval Filmus for his helpful insights leading us to find useful literature on this topic. Thanks are also due to Flavia Bonomo for relevant discussions.

References

Abeliuk, A.; Berbeglia, G.; Cebrian, M.; and Van Hentenryck, P. 2016. Assortment optimization under a multinomial logit model with position bias and social influence. *4OR* 14(1):57–75.

Anderson, S. P.; Depalma, A.; and Thisse, J.-F. 1992. *Discrete Choice Theory of Product Differentiation*. The MIT Press.

Bandyapadhyay, S. 2014. A variant of the maximum weight independent set problem. *CoRR* abs/1409.0173.

Ben-Akiva, M., and Lerman, S. 1985. *Discrete Choice Analysis: Theory and Application to Travel Demand*. MIT Press series in transportation studies. MIT Press.

Brady, R. L., and Rehbeck, J. 2016. Menu-dependent stochastic feasibility. *Econometrica* 84(3):1203–1223.

Bront, J. J. M.; Méndez-Díaz, I.; and Vulcano, G. 2009. A column generation algorithm for choice-based network revenue management. *Operations Research* 57(3):769–784.

Camion, P. 1965. Characterization of totally unimodular matrices. *Proceedings of the American Mathematical Society* 16(5):1068–1073.

Catalano, M.; Lo Casto, B.; and Migliore, M. 2008. Car sharing demand estimation and urban transport demand modelling using stated preference techniques.

Daly, A., and Zachary, S. 1978. Improved multiple choice models. *Determinants of travel choice* 335:357.

Davis, J.; Gallego, G.; and Topaloglu, H. 2013. Assortment planning under the multinomial logit model with totally unimodular constraint structures. *Department of IEOR, Columbia University. Available at http://www. columbia. edu/ gmg2/logit_const. pdf.*

Debreu, G. 1960. Review of R. D. Luce Individual Choice Behavior. *American Economic Review* 50:186–188.

Echenique, F., and Saito, K. 2015. General Luce Model. *California Institute of Technology SS Working Paper 1407.*

Echenique, F.; Saito, K.; and Tserenjigmid, G. 2013. The perception-adjusted luce model. *California Institute of Technology SS Working Paper 1379.*

Feldman, J., and Topaloglu, H. 2015. Bounding optimal expected revenues for assortment optimization under mixtures of multinomial logits. *Production and Operations Management* 24(10):1598–1620.

Ford, L. R. 1957. Solution of a ranking problem from binary comparisons. *The American Mathematical Monthly* 64(8):28–33.

Fudenberg, D., and Strzalecki, T. 2015. Dynamic logit with choice aversion. *Econometrica* 83(2):651–691.

Gallego, G.; Ratliff, R.; and Shebalov, S. 2011. A general attraction model and an efficient formulation for the network revenue management problem. *Technical Report, Columbia University, New York, NY.*

Gensch, D. H. 1985. Empirically testing a disaggregate choice model for segments. *Journal of Marketing Research* 22(4):462–467.

Guadagni, P. M., and Little, J. D. C. 1983. A logit model of brand choice calibrated on scanner data. *Marketing Science* 2(3):203–238.

Gul, F.; Natenzon, P.; and Pesendorfer, W. 2014. Random choice as behavioral optimization. *Econometrica* 82(5):1873–1912.

Hausman, J., and McFadden, D. 1984. Specification tests for the multinomial logit model. *Econometrica* 52(5):1219–40.

Jagabathula, S., and Vulcano, G. 2015. A model to estimate individual preferences using panel data. *Available at SSRN 2560994.*

Jagabathula, S. 2014. Assortment optimization under general choice. *Available at SSRN* 1598–1620.

Kalra, T.; Mathew, R.; Pal, S. P.; and Pandey, V. 2017. *Maximum Weighted Independent Sets with a Budget.* Cham: Springer International Publishing. 254–266.

Luce, R. D. 1959. *Individual Choice Behavior.* John Wiley and Sons.

Manzini, P., and Mariotti, M. 2014. Stochastic choice and consideration sets. *Econometrica* 82(3):1153–1176.

Masatlioglu, Y.; Nakajima, D.; and Ozbay, E. Y. 2012. Revealed attention. *American Economic Review* 102(5):2183–2205.

McFadden, D. 1974. Conditional logit analysis of qualitative choice behavior. *Frontiers in Econometrics* 105–142.

McFadden, D. 1978. Modeling the choice of residential location. *Transportation Research Record* (673).

Megiddo, N. 1979. Combinatorial optimization with rational objective functions. *Mathematics of Operations Research* 4(4):414–424.

Méndez-Díaz, I.; Miranda-Bront, J. J.; Vulcano, G.; and Zabala, P. 2014. A branch-and-cut algorithm for the latent-class logit assortment problem. *Discrete Applied Mathematics* 164, Part 1:246 – 263. Combinatorial Optimization.

Möhring, R. H. 1985. *Graphs and Order: The Role of Graphs in the Theory of Ordered Sets and Its Applications.* Dordrecht: Springer Netherlands. chapter Algorithmic Aspects of Comparability Graphs and Interval Graphs, 41–101.

Negahban, S.; Oh, S.; and Shah, D. 2012. Iterative ranking from pair-wise comparisons. In *Advances in Neural Information Processing Systems*, 2474–2482.

Rieskamp, J.; Busemeyer, J. R.; and Mellers, B. A. 2006. Extending the bounds of rationality: Evidence and theories of preferential choice. *Journal of Economic Literature* 44(3):631–661.

Rusmevichientong, P., and Topaloglu, H. 2012. Robust assortment optimization in revenue management under the multinomial logit choice model. *Operations Research* 60(4):865–882.

Rusmevichientong, P.; Shmoys, D.; Tong, C.; and Topaloglu, H. 2014. Assortment optimization under the multinomial logit model with random choice parameters. *Production and Operations Management* 23(11):2023–2039.

Rusmevichientong, P.; Shen, Z.-J. M.; and Shmoys, D. B. 2010. Dynamic assortment optimization with a multinomial logit choice model and capacity constraint. *Operations research* 58(6):1666–1680.

Shum, H., and Trotter, L. 1996. Cardinality-restricted chains and antichains in partially ordered sets. *Discrete Applied Mathematics* 65(1):421 – 439.

Talluri, K., and Van Ryzin, G. 2004. Revenue management under a general discrete choice model of consumer behavior. *Management Science* 50(1):15–33.

Train, K. 2003. *Discrete Choice Methods with Simulation.* SUNY-Oswego, Department of Economics.

Williams, H. C. 1977. On the formation of travel demand models and economic evaluation measures of user benefit. *Environment and planning A* 9(3):285–344.