Can Adaptive Conjoint Analysis perform in a Preference Logic Framework?[1]

Adrian Giurca, Ingo Schmitt and Daniel Baier
{giurca, schmitt, daniel.baier }@tu-cottbus.de

Abstract. Research on conjoint analysis/preference aggregation/social choice aggregation is performed by more than forty years by various communities. However, many proposed mathematical models understand preferences as irreflexive, transitive and statical relations while there is human psychology research work questioning these properties as being not enough motivated. This works propose to position the conjoint analysis inside a logical framework allowing for non-transitive and globally inconsistent preferences. Using a preference logics one can define a logic-based utility allowing to obtain an aggregate semantics of the collective choice.

1 Introduction and Motivation

Conjoint Analysis (CA) in marketing research was introduced forty years ago [26] being influenced by economics ([36], [35]) and mathematical psychology ([39], [40], [7]). While the beginning was devoted mostly to understand how individuals evaluate products/services and form preferences (see, [26], [34], [43] and possibly others), in the last thirty years the CA literature focused more on predicting behavioral outcomes by using statistical methods and techniques ([8]) and this resulted in a widespread variation in CA practice. Recently, applications in innovation market were developed ([9]).

The traditional conjoint task is related to the rational economy model where agents tend to action towards maximizing their utilities.

While traditional models obtain significant results when processing *complete*, *transitive* and *acyclic* (consistent) preferences, many communities mention that such models are quite far from the real life. When asking people about thing they like, then they may not answer (*incompleteness*), or they may change their initial preferences due to reception of new information (*preference change*). In addition, while it seems that the preference system of one respondent must be non contradictory, when processing preferences from many respondents this assumption does not remain valid. Some of our previous work argued towards a logic-based model for conjoint analysis.

[1] This research is supported by (1)DFG Project SQ-System: Entwicklung von Konzepten für ein quantenlogikbasiertes Retrieval-Datenbank-Anfragesystem: Anfragesprache, interaktive Suchformulierungsowie effiziente Anfrageauswertung and (2) German Federal Ministry of Education and Research, ForMaT project (Forschung für den Markt im Team), Phase II, Innovationslabor: Multimediale Ähnlichkeitssuche zum Matchen, Typologisieren und Segmentieren

The research reported by [46] proposed a mathematical optimization approach by translating ratings into algebraic constraints, but such solution requires acyclicity and transitivity and not changing preferences. New debates on solution proposed by [46] were reported by [31] in the context of non-additive utility aggregations such as Choquet integral. However, none of these approaches consider non-transitive and/or cyclic preferences, [48].

[23] introduced a logic-based utility but the approach was limited by a number of assumptions such as *consistency* (acyclic preferences) *ignorance* (of neutral rated questions), *transitivity* and the restriction of using only 2 stimuli choice pair comparisons. Moreover, while it argued on the logical nature of the users ratings and rankings, it does not consider *preference change* and interview adaptation. Many of these restrictions were introduced by the method of computing the logic-based utility, basically adaptation of the weighted majority learning algorithm allowing only binary preference as input.

As discussed by [24], computing beliefs from ratings and rankings is much close to the mental expectations of respondents and identified three kinds of beliefs that can be obtained from question answers. The proposed framework considers consistent respondent belief sets but on belief sets aggregation there is no need to require consistency: moreover this is inline with the Arrow's impossibility theorem (see [5] and [6]).

Although traditional non-adaptive conjoint solutions require static, non-changing, preferences, when data collection is interactive one may experience *preference change*. Moreover, the actual online solutions on data collection show many cases when the data is collected over days and not by a standard survey in a contiguous manner. As such, respondents may remake-up their mind therefore change is frequently expected. Also, [24] pointed that may be useful to use weighted beliefs due to the imprecise nature of the user ratings. In addition, among other distinctions it was emphasized that while *individual beliefs are consistent* (no assumption of user irrationality), *collective beliefs may not be consistent.* In addition, while the AGM model [4] considered *consolidation* as a maintenance operation of removing some dispensable beliefs resulting in a consistent knowledge base, we would like to avoid such approach due to missing of motivated criteria with respect of belief elimination.

The goal of this paper is to argue on the opportunity to use a preference logics framework allowing non-transitivity and inconsistency in preference data.

2 Related Work

The classical model of computing an utility function is the additive linear model (see [8] for details). Basically, the overall utility is an additive linear combination on value scores adjusted with attribute scores and compensated with a constant depending on interview i.e.,

$$U(o_j) = \mu + \sum_{k=1}^{N} \sum_{l=1}^{n_i} \beta_{kl} \cdot x_{jkl}$$

where
$U(o_j)$ – is the total score on product profile o_j,
$\beta_{kl} = U_k(a_{kl})$ – is the user preference on value a_{kl} of attribute A_k, and
$x_{jkl} = \begin{cases} 1, \; if \; o_j.A_k = a_{kl} \\ 0, \; otherwise \end{cases}$,
μ is a calibration constant (mean preference value across all objects). Usually $U_k()$ is called part-utility function or part-worth function and its specification depends of the attribute type (categorical and quantitative).

In practice a conjoint study may contain both types of attributes. Significant examples of categorical attributes are brand names or verbal descriptions containing levels such as "high", "medium", "low" while quantitative attributes are the ones which are measurable on either an interval scale or a ratio scale (e.g., speed of a processor, size of a screen). While there were proposed many models to encode the part-worth functions, two models are representative:

1. the *vector model*, $U_k(a_{kl}) = w_k\theta_{kl}$, where w_k is the weight of attribute A_k, and θ_{kl} is the weight of the value $a_{kl} \in dom(A_k)$) and
2. the *ideal point model*, $U_k(a_{kl}) = w_k(\theta_{kl} - \theta_{k0})^2$, where θ_{k0} is the weight of the ideal value a_{k0} of attribute A_k.

In overall the standard conjoint problem reduces to find all β_{kl} and μ by using training data of user-rated utilities for a training object dataset.

2.1 Machine Learning Approaches

During the last thirty years, Machine Learning research developed very similar problems, offering either statistically-based or logic-based solutions. As in traditional conjoint analysis, the difficulty relates to the fact that the set of all possible behaviors given all possible inputs is too large to be covered by the set of observed examples (training data). Hence the learner must generalize from the training data. Learning from examples towards forecasting the future behavior is one large field of research.

2.1.1 Support Vector Machines

Support Vector Machines, [10], [47] was proposed as a classification methodology by machine learning community. Basically, the standard model takes a set of input data and, classify each given input as being part of one of two possible categories (such as "like" and "unlike"). There is research proposing to use this model on conjoint analysis too (e.g., [16]).

The main assumptions of this method are: (a) there is preference data for a set of objects $\mathcal{O}$ and (b) the utility function is linear. Each preference data (e.g., $o_1 \preceq o_2$) is translated into an inequality between corresponding utilities of the corresponding objects ($u(o_1) \leq u(o_2)$). The method then involves minimizing the sum of errors for the inequalities and the sum of of the squares of the weights in the utility function.

As usual, each attribute value $a_{ij} \in dom(A_i), \; i = 1, ...n$ has weight θ_{ij}. We denote $\overline{\theta}_j^{(k)}$ the weights vector corresponding to the k-th object $o_j^{(k)}$. The goal is to estimate the individual partworths $w = (w_1, ..., w_n)$ considering a linear utility function (e.g., the vector model) $U(o) = \overline{w} \cdot \overline{\theta}$ for each $\overline{\theta}$ corresponding to an object $o \in \mathcal{O}$.

We encode preference data by respondent interviews: at the k-th question we show a subset $\mathcal{O}_k = \{o_1^{(k)}, ..., o_{n_k}^{(k)}\} \subset \mathcal{O}$ asking the respondent to choose one object as "the most liked". Without loosing the generality (via reordering) we can assume that the respondent choose first object as the preferred one. This choice is encoded as the set of constraints, $\overline{w}(\theta_1^{(k)} - \theta_i^{(k)}) \leq 0, i = 2, ..., n_k$, and reduce the conjoint problem to a classification problem. [16] proposes to train a L_2-soft margin classifier only with positive examples obtained from respondent ratings, using a with a hyperplane through the origin and modeling the answering noise with dummy variables $\varepsilon_i^{(k)}$. It trains one algorithm per respondent to get individual vector weights $\overline{w}^{(p)}$ for each respondent p and then to compute individual partworths by calibration with the aggregated partworths i.e. $\widetilde{w} = \frac{1}{|\mathcal{P}|}\sum_{p\in\mathcal{P}} \overline{w}^{(p)}$ and then $\overline{w}_*^{(p)} = \frac{\overline{w}^{(p)} + \widetilde{w}}{2}$. The training conditions are:

$$\begin{cases} Minimize : \overline{w}^2 + C \sum_{p_k\in\mathcal{P}} \sum_{i=2}^{n_k} (\varepsilon_i^{(k)})^2 \\ suchthat : \overline{w}(\theta_1^{(k)} - \theta_i^{(k)}) \leq 1 - \varepsilon_i^{(k)} \end{cases}$$

where C is a constant depending on the respondents set.

2.1.2 Learning from Preferences

Recall the learning problem similar with most of conjoint analysis tasks:

> Given a (very large) set of objects (each object represented as a set of attribute-value pairs), and a set of evaluation instances (each object is evaluated by experts obtaining a score, typically a real number) find a learning algorithm being able to evaluate any subset of the initial set of objects being compliant with expert evaluations.

As learning algorithms use evaluated training data it looks straightforward to input the learner with a database of examples in which the human expert has entered scores for each possible choice. However, similar with traditional conjoint analysis, there are two critical issues of this approach: (a) many domains have very large set of possible objects therefore is would be a tremendously time consuming for the expert to create the complete evaluation rank. Moreover, the training dataset must also contain enough "bad" alternatives otherwise the expert will be tempted to produce only high scores for everything and as such, to obtain a rank which is not useful; (b) in many cases experts do not think in terms of absolute scoring functions therefore will be very difficult, sometimes impossible, to create training data containing absolute scores. These reasons yields many researchers to consider pair comparisons rather than scoring individual alternatives(there is a large literature concerning the way users create preferences. The reader may consider [37], [12], [40], [17] and probably many other). Preference learning was pioneered by [53] and continued by [55], [33], [20] and possibly others. Basically, given a set of (partial) profiles and a preference function of these profiles we want be able to train a computer program to classify new (so far unseen) profiles by assigning a correct rank to each profile. The ratio of correctly classified data points is called the accuracy of the system.

As such conjoint analysis is similar with a learning task: *learning utility functions from respondent preferences.* The conjoint problem can be seen as learning to rank a set of objects by combining a given collection of initial rankings or preference functions. In machine learning community this problem of combining preferences arises in several applications, such as that of combining the results of different search engines, or the collaborative filtering problem. During the last 20 years a number of algorithm were developed: a pioneering algorithm is described in [14] and [15] as an extension of the early work reported by [38]. Advances in learning from preferences were reported by [19], [20], and [30]. As described by [20], the task of learning object preferences is:

> Let $\mathcal{O} = \{(a_1, ..., a_n)|a_i \in dom(A_i)\}$ be the set of all possible product representations and let $\mathcal{S} = \{o_1, ..., o_n\} \subseteq \mathcal{O}$ be a set of training objects (aka full profiles, product representations). Let $\mathcal{P}$ be a set of respondents and $\{P_{\mathcal{S},p} : \mathcal{S} \times \mathcal{S} \longrightarrow \{0,1\}|p \in \mathcal{P}\}$ the set of pairwise preferences on training data. Learn a utility function $U : \mathcal{O} \longrightarrow \mathbb{R}$ that ranks any subset of $\mathcal{O}$.

Notable, while conjoint analysis typically assume a linear utility function (see details by [8]), learning from preferences does not require utility linearity but many strategies on learning from preferences still assume linear combinations as potential ranking functions. A significant solution introduced by [14] and improved in [15] considers learning a global preference as a weighted linear combination of all respondent preferences, and then derive a final ordering which is maximal consistent with this preference. Other research ([53], [30]) uses a different strategy, specifically direct learning of the utility function directly from the respondent preferences. [53] introduces a two-state symmetric neural network architecture that can be trained with representations of states and a training signal (corresponding to the user preferences) indicated the preferred state. Subsequent works on this solution were reported by [55], [29], [33], and [27].

2.1.3 Logic-based Approaches

A logic-based approach was proposed by [46]by replacing the utility function with a logical formula best fulfilling a set of algebraic constraints derived from preference processing. They use Commuting Quantum Query Language (CQQL, [45]) a logical language based on combinations between Boolean conditions and proximity/similarity conditions over specialized variants of logical operators producing weighted formulas. The problem is formulated as below:

> Let $\mathcal{O} = \{(a_1, ..., a_n)|a_i \in dom(A_i)\}$ be the set of all possible object representations and $\mathcal{S} \subseteq \mathcal{O}$ a set of training objects. $\preceq$ denotes the preference relation on training data $\mathcal{S}$. Find a weighted full DNF CQQL formula $U = \bigvee_j w_j m_j$ (m_j is the j-th minterm and $w_j \in [0,1]$ its weight) such that U best fulfills the user preferences i.e. when CQQL evaluation is performed over objects in $\mathcal{O}$ then the obtained rank is consistent with user initial preferences.

If $o_{i_2} \preceq o_{i_1}$ then the following constraint is considered

$$eval_{CQQL}(U, o_{i_1}) - eval_{CQQL}(U, o_{i_2}) \geq 0$$

Because CQQL evaluation has simple arithmetic rules for formula evaluation, from the computational point of view the problem reduces to a linear optimization: $Maximize$: $\sum_{o_{i_2} \preceq o_{i_1}} (eval_{CQQL}(U, o_{i_1}) - eval_{CQQL}(U, o_{i_2}))$ under the above described constraints. The readers may consider [46] for details on problem solving strategies (such as simplex computations, feasible and unfeasible states, solutions to avoid overfitting and more.)

Automated extraction of rules from evidences was largely discussed by connectionist learning community (early work by [41], pioneered by [21] and subsequently discussed by [51], [25], [52], [11], [49], and possibly others) under the umbrella of a much general task:

> How can we extract models from the training data in an automated manner and use these models as the basis of an autonomous rational agent in the given domain.

One of the most important features of such an approach is that it combines the computational advantages of connectionist models with the qualitative knowledge representation proposed by the AI community.

It is obvious that a solution of this problem must consider two stages: (1) Learning the model and (2)Performing inference using this model. This work follows only the first stage of the problem – if there is a learned ruleset then there are many opportunities to perform inference according with various semantics (crisp, probabilistic, fuzzy and so on) and a discussion of appropriateness of each of them should be large.

Inside a rule framework the conjoint problem is to find out a set of rules that best model the respondent preferences. One can consider learning of various kinds of rules (possibly weighted), each of them supporting various semantics including probabilistic models [42], incomplete/imprecise information, [54], plausibility-based models [18], [22] or quantum logic semantics [45]:

1. Simple rules (propositional rules):
$$[(\neg)A_{i_1} \wedge ..., \wedge(\neg)A_{i_k} \rightsquigarrow A_{i_{k+1}}]$$
where $(\neg)A$ denotes a possibly negated attribute;
2. Positive attribute-value rules:
$$[A_{i_1} \simeq v_{i_1} \wedge ..., \wedge A_{i_k} \simeq v_{i_k} \rightsquigarrow A_{i_{k+1}} \simeq v_{i_{k+1}}]$$
where $v_{i_j} \in dom(A_{i_j})$, $A_{i_j} \simeq v_{i_j}$ means that A_{i_j} takes a value around v_{i_j} (The reader should notice that $\simeq$ includes ordinal values, e.g., $A_{i_j} = v_{i_j}$);

3. Attribute-value rules with negation:

$$[(\neg)A_{i_1} \simeq v_{i_1} \wedge ..., \wedge(\neg)A_{i_k} \simeq v_{i_k} \rightsquigarrow A_{i_{k+1}} \simeq v_{i_{k+1}}]$$

where $\neg A_{i_j} = v_{i_j}$ means $A_{i_j} \neq v_{i_j}$;

4. General attribute-value rules:

$$[(\neg)A_{i_1} \simeq v_{i_1} \wedge ..., \wedge(\neg)A_{i_k} \simeq v_{i_k} \rightsquigarrow (\neg)A_{i_{k+1}} \simeq v_{i_{k+1}}]$$

The first three kinds of rules were largely addressed by data mining community when learning association rules. Researchers developed different kinds of association rules: Boolean (crisp) association rules, quantitative association rules, fuzzy association rules. Association rules were pioneered by [44] and then established by [2], and [3]). Standard association rules consider two measures of interestingness: *support* and *confidence* although other models may add two more: *lift* and *conviction* or adopt non-standard ones, [32]. Learning association rules is usually performed under both a user-specified minimum support and a user-specified minimum confidence requirements.

There were developed many algorithms starting with the most known one, Apriori ([3]) and continuing with many others (Eclat, FP-growth and so on.) A significant step is the Assoc algorithm [28] which enables mining for generalized association rules (including negation i.e. attribute-value rules with negation) and does not restrict for minimum support and confidence.

However, on our knowledge, none of this research considering the conjoint analysis task: basically the training data set for learning association rules does not distinguish various users. All the data is uniform (mostly, it comes from e-commerce transactions) and it may refer to one user (such as in recommender systems, [1]) or to many but not considering distinct training data for each of them, therefore the conjoint task is somehow hidden. In addition the conjoint analysis problem in the context of learning association rules does not directly performs from preferences: using transactional data as input, there should be some algorithm computing binary preferences.

The first kind of rules were considered, in context of adaptive conjoint analysis, by [23] in conjunction with weighted CQQL (see [45] for language description), an extension of the relational calculus using quantum logic paradigm which defines *metric*(or *similarity*) predicates, weighted conjunction ($\wedge_{\theta_1,\theta_2}$), weighted disjunction ($\vee_{\theta_1,\theta_2}$) and quantum negation. Clearly (as explained by [25] and [52]) there is a need for both a preference measure to rank the rules and a learning algorithm which uses the preference measure to find the best k rules. The work reported by [23] describes a heuristic and learning approach to use the respondent preferences on stimuli to compute a rule preference relation (called minterm preference because the rules were learned as weighted minterms of the CQQL full disjunctive normal form) and then use a learning algorithm to compute a ranking on the minterms set.

3 Conjoint Analysis using Preference Logics

This section introduces a logical framework allowing (a) encoding of preferences as choice formulas, (b) defining a logic-based utility inside a preference logic to allow creation of collective beliefs and (c) performing rule extraction and explanation and formal interpretation.

3.1 Preference Logics

We follow the approach defined by [50] on preference logic introduced as a special case of logic by defining a preference relation between the interpretations of the underlining logic as we consider this approach being simple and powerful. Below we recall some of the [50] results.

Let $\mathcal{L}$ be a standard logic and $\sqsubset$ a strict partial order on interpretations (we say $\mathcal{I}_2$ is preferred to $\mathcal{I}_1$ and denote $\mathcal{I}_1 \sqsubset \mathcal{I}_2$). Then, $\mathcal{L}_\sqsubset = (\mathcal{L}, \sqsubset)$ is a new logic, a preference logic. The basic artifacts such as satisfaction, validity and entailment are defined by [50]. Recall that while the standard logics are monotonic[2]. Recall the definitions of satisfiability, validity and entailment:

Definition 1 ([50])
Let $F, G \in \mathcal{L}$. Let $\mathcal{I}$ be an interpretation.
$\mathcal{I}$ preferentially satisfies F *(denoted $\mathcal{I} \models_\sqsubset F$) if $\mathcal{I} \models F$ and there is no $\mathcal{I}'$ such that $\mathcal{I} \sqsubset \mathcal{I}'$ and $\mathcal{I}' \models F$. As usual, $\mathcal{I}$ is called the* model *of F.*
F preferentially entails G *(denoted $F \models_\sqsubset G$) if*

$$\forall \mathcal{I}, \mathcal{I} \models_\sqsubset F \Rightarrow \mathcal{I} \models_\sqsubset G$$

That is the preferred models of G are also preferred models of F.

As described by [50], $\mathcal{L}_\sqsubset$ is a non-monotonic logic because there may be formulas $F, G \in \mathcal{L}_\sqsubset$ such that both $F \models_\sqsubset G$ and $F \models_\sqsubset \neg G$. Moreover, it is not necessary that F is inconsistent, it is just sufficient that F do not have preferred models.

A significant case of preference logics was introduced by [13] under the name of choice logic. Basically, choice logic defines the ordered disjunction (denoted $\times$) as a special kind of standard disjunction ($\vee$) as such introducing a preference relation between the interpretations and models. The ordered disjunction has the same models as regular disjunction but there is a preference relation between these models. For example, if $A \times B$ is a disjunction between two atoms. Then $\mathcal{I}_1 = \{A\}$, $\mathcal{I}_2 = \{A, B\}$ and $\mathcal{I}_3 = \{B\}$ are its models. Then $\mathcal{I}_3 \sqsubset \mathcal{I}_2$ and $\mathcal{I}_3 \sqsubset \mathcal{I}_1$ meaning that $\mathcal{I}_1$ and $\mathcal{I}_2$ are preferred models.

Intuitively, as [13] reports, the ordered disjunction means that when $F_1 \times ...F_n$ we prefer models that first satisfies F_1 and if this is not possible then we prefer models satisfying F_2, and so on. Choice logic defines the degree of satisfaction for all logic formulas

Definition 2 ([13])
The optionality *of a formula (the number of choices to satisfy a formula) is $opt(A) = 1$ if A is an atom.*
$opt(\neg F) = 1$
$opt(F_1 \vee F_2) = \max(opt(F_1), opt(F_2))$
$opt(F_1 \wedge F_2) = \max(opt(F_1), opt(F_2))$
$opt(F_1 \times F_2) = opt(F_1) + opt(F_2)$

[13] defines the preference relation ($\sqsubset$) between models of logic formulas and consequently the entailment. It is shown that

[2] In the sense that if $F_1, F_2, F_3 \in \mathcal{L}$, if $F_1 \models F_3$ then $F_1 \wedge F_2 \models F_3$.

the entailment satisfies cautious monotony and cumulative transitivity:

Proposition 1 ([13])
Let S be a set of choice logic formulas and A, B be classical formulas.
$S \models_{\sqsubset} A$ *and* $S \models_{\sqsubset} B \Rightarrow S \cup \{A\} \models_{\sqsubset} B$
$S \models_{\sqsubset} A$ *and* $S \cup \{A\} \models_{\sqsubset} B \Rightarrow S \models_{\sqsubset} B$

From the computational point of view, choice logic can be translated to stratified knowledge bases.

4 Modeling Conjoint Analysis

Conjoint analysis collects preferences from user interviews using a variety of question types but the most used ones are trade-off matrices and pair-comparisons. A trade-off matrix ([34]) asks a respondent to consider a pair of attributes. It displays all combinations of values for those attributes, asking the respondents to provide a ranking for the combinations. The Table 1 show an example of a trade-off matrix related to attributes *OperatingSystem* and *Battery life*. While trade-off

	12 hours	6 hours	4 hours	2 hours
Android	1	2	7	5
WinPhone	3	4	6	11
other OS	8	9	10	12

Table 1. A trade-off matrix with respondent ranking

matrix are quite efficient on ranking binary stimuli, trade-off matrices cannot be used if we consider stimuli with more than two attributes. A solution to these limitations is to use pair comparisons. Pair comparisons are seen as choice questions

Left side	OR	Right side
Android AND $\geq$ 500EUR, ...	Left	Windows Phone,... AND $\leq$ 3.5" screen, ...
$\geq$ 4" screen,... AND Battery life 6h	Neutral	And,... AND WIFI, ...
$\geq$ 4" screen,... AND other OS	Left	Battery life 10h,... AND no WIFI, ...

Table 2. Pair Comparisons and Ratings

evaluated by favoring either "the left side" or "the right side" or "neutral".

4.1 Preferences as Choice Formulas

Let $A_1, ..., A_n$ be a set of attributes (unary predicates) with $dom(A_i)$ the domain of values. Let $\mathcal{O} = \{(a_1, ..., a_n)|a_i \in dom(A_i)\}$ be the set of all possible product representations. The choice logic ordered disjunction operator makes this logic suitable candidate to encode user ratings as choice formulas. The trade-off matrices introduces a rank between choices e.g., the matrix from Table 1 say that $OS(\text{"}Android\text{"}) \wedge Battery(\text{"}12h\text{"})$ is preferred to $OS(\text{"}Android) \wedge Battery(\text{"}6h\text{"})$ as well as $OS(\text{"}WinPhone\text{"}) \wedge Battery(\text{"}12h\text{"})$ is preferred to $OS(\text{"}Android) \wedge Battery(\text{"}4h\text{"})$ and so on.

Definition 3 (Mapping trade-off matrices)
Let a trade-off matrix based on predicates A_1 and A_2.
If $A_1(u) \wedge A_2(v)$ is preferred to $A_1(u') \wedge A_2(v')$ then this preference is encoded into the choice formula:

$$A_1(u) \wedge A_2(v) \times A_1(u') \wedge A_2(v')$$

that is preferring models that, if possible first satisfy $A_1(u) \wedge A_2(v)$[3].

Definition 4 (Mapping pair comparisons)
Let q be the pair comparison
$q = A(a)$ *and* $B(b)$ *OR* $C(c)$ *and* $D(d)$.
If the left side is preferred then this preference is encoded into the choice formula:

$$A(a) \wedge B(b) \times C(c) \wedge D(d)$$

If q is rated neutral then this preference is encoded into the formula:

$$A(a) \wedge B(b) \vee C(c) \wedge D(d)$$

Similarly, if the right side is preferred then this preference is encoded into the choice formula:

$$C(c) \wedge D(d) \times A(a) \wedge B(b)$$

4.2 Towards Logic-based Conjoint Analysis

Let $\mathcal{A} = \{A_1, ..., A_n\}$ be a set of unary predicates with $dom(A_i)$ the domain of values. Let $\mathcal{O} = \{(a_1, ..., a_n)|a_i \in dom(A_i)\}$ be the set of all possible product representations.

Definition 5 (Normal Form)
A full ordered disjunctive normal form *(ODNF) over choice logic defined by the language $\mathcal{A}$ is a formula*

$$U = \times_j (L_1(l_1^j) \wedge ... \wedge L_n(l_n^j))$$

where $L_k^(l_k^j)$ is a literal corresponding to the predicate A_k (either $A_k(l_k^j)$ or $\neg A_k(l_k^j)$) and $l_k^j \in dom(A_k)$.

Let $\mathcal{C}$ the set of all choice formulas derived from user preferences. Then, the generic conjoint analysis task is described as below:

> Find $U = \times_j (L_1(l_1^j) \wedge ... \wedge L_n(l_n^j))$ such that U best fulfills the user preferences i.e. there is a maximal set of constraints $\mathcal{C}' \subseteq \mathcal{C}$ such that $U \models_{\sqsubset} C$ for all $C \in \mathcal{C}'$.

Of course, the economics community does not really need the complete DNF but, most of the cases only a subset of the ODNF (the most important clauses). In addition, sometimes the constraints may come weighted (using some weight $w \in (0, 1]$) and then the concept of maximal set can be replaced by a subset of constraints with a sum of weights greater than a specified threshold.

[3] This corresponds completely to the psychological meaning of trade-off matrices where the respondent *does not reject* any of the alternatives

Rule extraction from a computed ODNF (or a subset) is straightforward as the experts like to understand the dependencies of a specific predicate value with respect of the remaining predicates. As such rules are obtained by transforming U to conjunctive normal form (CNF) and then deriving rules from each clause according with specific predicates as conclusions.

Let $\mathcal{R}$ be a the derived ruleset as described above. Then, all preferred models of $\mathcal{R}$ corresponds to preferred objects in $\mathcal{O}$.

As such we propose an updated process chain of adaptive logic-based conjoint analysis as depicted by Figure 1.

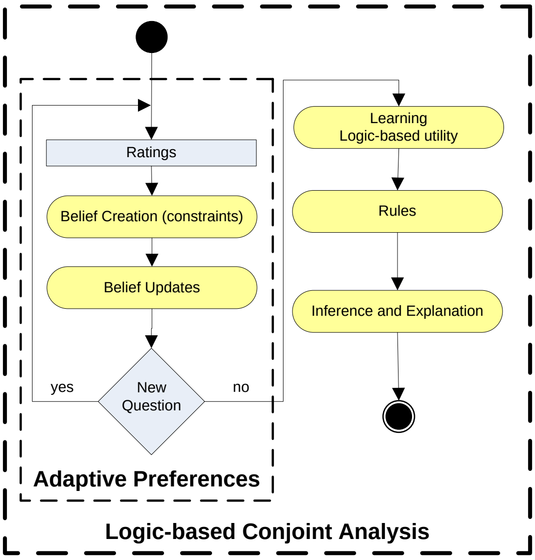

Figure 1. Logic-Based Adaptive Conjoint Analysis Chain

5 Conclusion

We proposed a model of logic-based conjoint analysis by considering encoding respondent preferences as beliefs (as such allowing belief change) and encoding this beliefs to choice formulas. While the individual beliefs translates into consistent constraints set the collective beliefs (all constraints collected from all respondents) may not be a consistent set. The Table 3 describes the kind of preferences used by the analyzed models. As seen the proposed approach is useful when the model intends to capture phycological phenomena such as change or irrationality (inconsistency) as well as when formal explanations of decisions need to be computed. This work is at its beginnings: beside fine tuning and debugging, obtaining feasible algorithms to compute the logic-based utility is a mandatory next step. Analyzing such algorithms may open discussion on improvements of the preference logic too as traditional processing of pair comparisons also consider Likert scales as rating methods. In addition, a close look on the necessary belief framework (a discussion was started by [24]) is necessary.

Aggregation Models	Require Irreflexive	Require Transitive	Allow Indifference	Static Preference
CA (econ.)	yes	yes	yes	yes
SVM	yes	yes	no	yes
Preference Learning	yes	yes	no	yes
Rule Learning	yes	yes	no	yes
Preference Logic	yes	**no**	yes	**no** (belief rev)

Table 3. Conjoint Analysis Preference Requirements

REFERENCES

[1] G. Adomavicius, and A. Tuzhilin, Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions, IEEE Transactions on Knowledge and Data Engineering, Vol. 17, No. 6, June 2005, pp. 734-749.

[2] R. Agrawal, T. Imielinski, and A.N. Swami. Mining association rules between sets of items in large databases. In P. Buneman and S. Jajodia (Eds.), Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data, Washington, D.C., pp. 207-216, May 26-28, 1993.

[3] R. Agrawal, and R. Srikant. Fast algorithms for mining association rules. In J. B. Bocca, M. Jarke, and C. Zaniolo (Eds.), Proc. 20th Int. Conf. Very Large Data Bases, (VLDB), pp. 487-499, Morgan Kaufmann, 1994.

[4] C.E. Alchourron, P. Gärdenfors and D. Makinson. On the Logic of Theory Change: Partial Meet Contraction and Revision Functions. Journal of Symbolic Logic, 50: 510-530, 1985.

[5] K.J. Arrow. A Difficulty in the Concept of Social Welfare. Journal of Political Economy 58(4) (August, 1950), pp. 328-346.

[6] K. J. Arrow. Social Choice and Individual Values. 2nd ed., 1963.

[7] N.H. Anderson. Foundations of information integration theory. Academic Press, 1981.

[8] D. Baier and M. Brusch (Eds.) Conjointanalyse, Methoden - Anwendungen - Praxisbeispiele, Springer, Berlin, 2009.

[9] D. Baier. Conjoint Measurement in der Innovationsmarktforschung, in: Baaken, Thomas; Höft, Uwe; Kesting, Tobias (Hrsg.), Marketing für Innovationen - Wie innovative Unternehmen die Bedürfnisse ihrer Kunden erfüllen, Harland Media, Münster, ISBN-13 978-3-938363-42-3.

[10] B. E. Boser, I. M. Guyon, and V. N. Vapnik. A training algorithm for optimal margin classifiers. In Proc. 5th Annu. Workshop on Comput. Learning Theory, 1992.

[11] O. Boz. Knowledge Integration and Rule Extraction. Neural Networks, University of Leigh, 1995.

[12] R.A. Bradley and M.E. Terry. Rank analysis of incomplete block designs: the method of paired comparisons. Biometrika, 39 (3-4), 1952, pp.324-345.

[13] G. Brewka, S. Benferhat and D. Le Berre. Qualitative Choice Logic. Proceedings of the Eights International Conference on Principles and Knowledge Representation and Reasoning (KR-02), Toulouse, France, April 22-25, 2002, pp.158-169.

[14] W. Cohen, R.E. Schapire and Y. Singer. Learning to Order Things. Advances in Neural Information Processing Systems 10, Morgan Kaufmann, 1998.

[15] W. Cohen, R.E. Schapire and Y. Singer. Learning to Order Things. Journal of Artificial Intelligence Research 10, pp. 213-270, 1999.

[16] T. Evgeniou, C. Boussios, and G. Zacharia. Generalized robust conjoint estimation. Marketing Science, 25, 2005.

[17] J. Eliashberg. Consumer Preference Judgments: An Exposition with Empirical Applications. Management Science, 26, 1, (January), 1980, pp.60-77.

[18] N. Friedman, and J.Y. Halpern. Plausibility measures and default reasoning. Journal of the ACM, 48, 2001, pp.648-685.

[19] J. Fürnkranz and E. Hüllermeier. Pairwise preference learning and ranking. Procs. of the 14th European Conference on Machine Learning (ECML-03), LNAI 2837, 2003, pp.145-156, Springer Verlag, 2003.

[20] J. Fürnkranz and E. Hüllermeier. Preference learning. Künstliche Intelligenz, 19(1), pp. 60-61, 2005.

[21] S.I. Galant. Connectionist Expert Systems. Communications of ACM, 31, 1988, pp.152-169.

[22] A. Giurca. A Logic with Plausibility. Annales of Craiova University, Mathematics and Computer Science Series, XXVII, pp.105-115, 2000.

[23] A. Giurca, I. Schmitt, and D. Baier. Performing Conjoint Analysis within a Logic-based Framework. Proc of IEEE Federated Conference on Computer Science and Information Systems, (FedCSIS2011), Szczecin, Poland, 18-21 September, 2011.

[24] A. Giurca, I. Schmitt, and D. Baier. Adaptive Conjoint Analysis. Training Data: Knowledge or Beliefs? A Logical Perspective of Preferences as Beliefs, KAM'2012 - 18th Conference on Knowledge Acquisition and Management, at FEDCSIS 2012, Wroclaw, Poland.

[25] R. M. Goodman, C. M. Higgins, J. W. Miller, and P. Smyth. Rule-Based Neural Networks for Classification and Probability Estimation. Neural Computation 4(6), pp.781-804, 1992.

[26] P. E. Green and V. Rao. Conjoint measurement for quantifying judgmental data. Journal of Marketing Research, 8, 1971, pp.355-363.

[27] P. Haddawy, V. Ha, A. Restificar, B. Geisler, and J. Miyamoto. Preference elicitation via theory refinement. Journal of Machine Learning Research, 4, pp. 317-337, 2003.

[28] P. Hájek. The new version of the GUHA procedure ASSOC, COMPSTAT 1984, pp.360-365.

[29] Ralf Herbrich, Thore Graepel, Peter Bollmann-Sdorra, and Klaus Obermayer. Supervised learning of preference relations. Procs. des Fachgruppentreffens Maschinelles Lernen (FGML-98), 1998, pp. 43-47.

[30] E. Hüllermeier, J. Fürnkranz, W. Cheng, and K. Brinker. Label ranking by learning pairwise preferences, Artificial Intelligence, Volume 172, Issues 16-17, pp. 1897-1916, 2008.

[31] E. Hüllermeier and I. Schmitt. Non-Additive Utility Functions: Choquet Integral versus Weighted DNF Formulas, The 4th Japanese-German Symposium on Classification (JGSC2012),March 9-10, 2012, Kyoto, Japan.

[32] I. Iancu, M. Gabroveanu and A. Giurca. A Pair of Confidence Measures for Association Rules, 30th Annual Conference of the German Classification Society, GfKl2006, March 8-10, 2006, Berlin, Germany.

[33] T. Joachims. Optimizing search engines using clickthrough data. Procs. of the 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-02), pp. 133-142. ACM Press, 2002.

[34] R. M. Johnson. Tradeoff Analysis of Consumer Values. Journal of Marketing Research, 1974, pp. 121-127.

[35] R.L. Keeney and H. Raiffa. Decisions with multiple objectives: Preferences and value tradeoffs. Wiley Series in Probability and Mathematical Statistics. NY: John Wiley & Sons, 1976.

[36] K. Lancaster. A new approach to consumer theory. Journal of Political Economy, 74, 1966, pp.132-157.

[37] R. Likert. A Technique for the Measurement of Attitudes. Archives of Psychology 140, 1932, pp.1-55.

[38] N. Littlestone and M. Warmuth. The weighted majority algorithm. Information and Computation, 108 (2), 1994, pp. 212-261.

[39] R.D. Luce and J. W. Tukey. Simultaneous Conjoint Measurement: A New Type of Fundamental Measurement. Journal of Mathematical Psychology, 1, 1964, pp.1-27.

[40] R.D. Luce and P. Suppes. Preference, utility and subjective probability. in Luce, R.D., Bush, R.R., and Galanter, E. (Eds.), Handbook of Mathematical Psychology, III, New York: Wiley, 1965, pp.235-406.

[41] M. C. Mozer. RAMBOT: A Connectionist Expert System That Learns by Example. Tech. Report. California Univ., San Diego, La Jolla. Inst. for Cognitive Science, 1986.

[42] N. J. Nilsson. Probabilistic logic. Artificial Intelligence 28(1), pp.71-87, 1986.

[43] K.L. Norman and J.J. Louviere. Integration of attributes in public bus transportation: two modeling approaches. Journal of Applied Psychology, 59, 6, 1974, pp.753-758.

[44] G. Piatetsky-Shapiro. Discovery, analysis, and presentation of strong rules, in G. Piatetsky-Shapiro and W. J. Frawley (eds): Knowledge Discovery in Databases. AAAI/MIT Press, Cambridge, MA, 1991.

[45] I. Schmitt. QQL: A DB&IR Query Language. VLDB Journal, 17(1), pp.39-56, 2008.

[46] I. Schmitt, and D. Baier. Logic Based Conjoint Analysis using the Commuting Quantum Query Language, Proc. of Conference of the German Classification Society (GfKl2011), August 31 to September 2, 2011, Frankfurt am Main, Germany.

[47] B. Schölkopf and A. Smola. Learning with kernels. MIT Press, 2002.

[48] G.F. Schumm. Transitivity, Preference and Indifference. Philosophical Studies, 52: 435-437, 1987.

[49] J. Sima. Neural Expert System. Journal of Neural Networks, vol. 8, no. 2, pp. 261-271, 1995.

[50] Y. Shoham. Nonmonotonic Logics: meaning and utility. Proceedings of 10th IJCAI, pp.388-393, Milan, 1987.

[51] P. Smyth, and R. M. Goodman. An Information Theoretic Approach to Rule Induction from Databases. IEEE Trans. Knowl. Data Eng. 4(4), pp.301-316, 1992.

[52] R. Sun. Integrating rules and connectionism for robust commonsense reasoning. Hoboken, N.J: Wiley & Sons, 1994, ISBN 0-471-59324-9.

[53] G. Tesario. Connectionist learning of expert preferences by comparison training. Advances in Neural Information Processing Systems, 1, pp. 99-106, Morgan Kaufmann, 1989.

[54] G. Wagner. Logic Programming with Strong Negation and Inexact Predicates. Journal of Logic and Computation 1(6), pp. 835-859, 1991.

[55] J. Wang. Artificial neural networks versus natural neural networks:A connectionist paradigm for preference assessment. Decision Support Systems, 11, pp. 415-429, 1994.