Proceedings of the 37th Design Automation Conference
IDETC/CIE 2011: DAC-25
Aug 28-31, 2011, Washington, DC, USA

DETC2011-48749

GRANULARITY ENHANCEMENT OF EXTRACTED PREFERENTIAL PROBABILITIES FROM DESIGN TEAM DISCUSSION

Haifeng Ji
Yahoo! Inc.
Sunnyvale, CA, USA

Tomonori Honda
Department of Mechanical Engineering
Massachusetts Institute of Technology
Cambridge, MA, USA

Maria C. Yang
Department of Mechanical Engineering and Engineering Division
Massachusetts Institute of Technology
Cambridge, MA, USA

ABSTRACT

Preferences are a formal way to represent a designer's choices when assigning priorities for a set of possible design alternatives within the context of the design process. A design team's preferences can change over the life of project, and knowledge of this evolution can be useful for understanding a team's rationale as well as its confidence in a decision. This paper presents a "sliding window" approach (SPPT) to the extraction of preference related information from transcribed design team discussion. The approach suggested in this paper can assess design preferences over time with a finer granularity than a previous approach known as PPT, and removes perturbations that occur when there is little design team discussion. Both SPPT and PPT were applied to a discussion transcript. Results show good consistency among SPPT, PPT and survey results. SPPT is also able to detect more changes in design team preference.

INTRODUCTION

Design preferences are the priorities assigned to a set of design choices. In a design project, understanding the preferences of a designer can be critical for making good design decisions. Several methodologies in engineering design research have been established to aid in making design decisions, including Pugh charts [1], Quality Function Deployment (QFD) [2], Decision-based Design (DBD) [3], and the Method of Imprecision (MoI) [4, 5]. Such approaches generally require a designer to state his or her explicit, quantitative preferences. However, when such methods are employed by a team of designers rather than by an individual, it is sometimes difficult to obtain the preferences of a group due to a number of challenges, including the overhead of eliciting preferences in a workplace setting, appropriately balancing the trade-offs between various design choices, and effectively aggregating those preferences into a single group preference. Previous research [6] presented a probabilistic approach to harvesting preference-related information from a team of designers that avoided some of these elicitation and aggregation issues. The approach, known as Preferential Probabilities from Transcripts (PPT), involves treating the team as a single entity and obtaining preference-related information in a descriptive fashion. In Honda, et al. [7], PPT was expanded to include formal linguistic appraisal to help identify designers' attitudes towards design alternative and extract corresponding preferential probabilities. Note that it is not a prescriptive decision making method but a descriptive method which helps understand the decision making during concept selection. In both approaches, rather than explicitly elicit preference-related information, they implicitly extract it from transcripts of team's discussion. It does not attempt to model design trade-offs but instead draws on the back-and-forth deliberation between team members to provide a pattern of preference-related information over the life of a project. The approaches avoid aggregating group information. Rather than focus on individual comments in discussion, they treat the entire team's discussion as a single collection of all team member comments. For the sake of simplicity, we call both of the above approaches PPT in this paper. The main difference is the first one uses word occurrences while the second one uses formal linguistic appraisal values.

The strategy of PPT is to divide design deliberation into several time intervals, then study the evolution of preferential probabilities from one interval to another. One of the assumptions of PPT is that a design team makes no preference changes within an interval. This assumption leads to the limitation that the interval granularity may not be fine enough to extract the decision-making that occurs within each interval. On the other hand, if we reduce the interval size too much, there may not be a sufficient number of samples per interval for effective analysis. In this work, we suggest an improved approach based on a "sliding window" module called Sliding-window Preferential Probabilities from Transcripts (SPPT), which helps us to access more detailed preference information from the design process.

In this paper, the SPPT approach is tested on a three-person design team working on a design selection problem in the re-design of a coffee maker. Periodic surveys were administered in the design process to collect individual designers' preference information. These survey results were then translated into preferential probabilities and compared against those extracted from transcripts. Both SPPT and PPT were compared against the baseline PPS (Preferential Probabilities from Surveys) [8].

RELATED RESEARCH

In any given design project, multiple criteria may be considered in design selection, and if any of these criteria are at odds, designers may need to make trade-offs among them. Methods for formalizing design trade-offs include utility theory [9-14] originally developed by von Neumann and Morgenstern and the Method of Imprecision (MoI) [4, 15] proposed by Otto, Antonsson, and Wood. Much research in engineering design decision-making has focused on how to assess preferences. A rigorous way to quantitatively value preference is by using the lottery method [16], which scales alternatives from 0 to 1. Formal approaches for eliciting preferences in pair-wise comparisons include the Analytical Hierarchy Process (AHP) [17] and the fuzzy outranking method [18]. For multi-attribute design, aggregation functions [19] can be used to formally calculate overall preferences. There are also methods based on factorial design and statistics for eliciting preferences, such as conjoint analysis [20] and discrete choice analysis [21, 22].

These above methods either assume a preference value for each criterion, or require explicitly asking designers for their preferences. PPT [6] is an implicit way to extract preference-related information from design team discussion, and this study improves the previous PPT approach and tries to extract the preferences in a finer mode which would help us understand the design process better.

There is a rich literature on decision-making styles in groups and how the opinions of team members and team leaders may be aggregated qualitatively [23]. However, research on decision-making in engineering design has centered around formal methods for aggregation. Arrow's Impossibility Theorem [24, 25] states that there is no guarantee of consistency for aggregation of preferences in a group. Still, there are a number of ways individual rating may be combined. For aggregating individuals' preference ordering, Dym, Wood and Scott [26] describe a structured pairwise comparison chart (PCC) which produces identical results with those produced by a Borda count. In aggregating individual's preference ratings, Keeney assumes each group member's opinion is equally important and uses cardinal utility functions to accumulate group preferences [27]. Jabeur, Martel et al. [28], and See and Lewis [29] study unequal weights on the preferences of the group members. In this study, both PPT and SPPT take an approach that does not require knowledge of an individual's preferences in order to formulate weightings, nor does it require aggregation of individual preferences. Rather, the group is treated as a single entity that generates information about the group's preferences during discussion. These group preferences are then extracted directly from transcripts of team discussion without aggregation.

APPROACHES TO IMPROVING THE GRANULARITY OF PREFERENCE INFORMATION EXTRACTION

PPT Using Small Intervals

Preferential Probabilities from Transcripts (PPT) is a method for assessing preference-related information from unstructured group design discussion. This method assumes that designers use natural language during design discussion that reflects their design process and can thus provide insights into their rationale and preferences [30, 31]. Preferential Probabilities from Transcripts (PPT) extracts preferential probabilities from the language generated during design team discussion. A detailed explanation of the formulation of PPT can be found in [32] and will only be described at a high level in this paper. PPT approximates the likelihood a design alternative will be "most preferred" based on what a design team says during a discussion. The transcript of the design team's discussion is divided into time intervals that allow observation of how preferential probabilities change over a series of intervals. It is implicit in PPT that what design team members say to each other during a design discussion largely corresponds with what they think [33].

PPT employs two basic models: the Preference Transition Model, which describes a relationship between what designers think from one time interval to the next, and the Utterance-Preference Model, which describes a relationship between what designers say and think within the same time interval. These two models are extracted from transcripts of team discussion and then used to predict latent preference data.

One of the limitations of the original formulation of PPT is that it relied on word occurrence and did not consider the possible linguistic implications of those words on a team's preferential probabilities. In Honda et al. [7], a linguistic approach was applied to provide a more nuanced analysis of team transcripts. The model of language used in that analysis is the system of APPRAISAL which construes affect and interpersonal relations based on the theory of Systemic

Functional Linguistics (SFL) [34]. An appraisal is the representation through language of favorable and unfavorable attitudes towards specific subjects. APPRAISAL, defines five semantic resources: (1) Attitude, (2) Engagement, (3) Graduation, (4) Orientation and (5) Polarity [35].

In PPT with appraisal [7], a positive (++) or negative (--) utility is used to model whether a designers' appraisal toward an alternative is increasing or decreasing. When appraisal was applied to PPT, negative appraisals were converted into positive appraisals of non-preferred alternatives. The appraisal values from the linguistic analysis and coding replaced word occurrences in the original PPT approach.

The approach for deriving PPT (including PPT with appraisal) from the transcript follows 6 steps:

1. Analyze the word occurrences (or appraisal values) of all design alternatives and synonyms in a transcript of a design team's discussion.
2. Construct an initial *utterance-preference model* and a *preference transition model* with hidden parameters. The utterance-preference model describes the likelihood the most-preferred alternatives and the less-preferred alternatives will be uttered in a discussion. The preference transition model describes how likely a most-preferred alternative is to change in the next interval.
3. Estimate reasonable initial parameters for the utterance-preference model and the preference transition model.
4. Apply the two models to predict preferential probabilities.
5. Apply a traditional Expectation-Maximum (EM) algorithm [36] to re-estimate the parameters of the two models with the utterance data and predicted preference data.
6. Iterate on Steps 4 and 5 until the hidden parameters of the models converge. As the EM algorithm improves the likelihood of the occurrences of the utterance data at each iteration [37], parameters are guaranteed to converge.

In applying PPT (or PPT with appraisal), it is assumed that designers do not change their preferences for a set of design alternatives within one time interval. Furthermore, preferences can only be changed at the transitions between time intervals.

SPPT: Sliding Window Approach to Extracting Probabilistic Preferences from Transcripts

The following notations were used throughout the paper for the sake of simplicity.

W: window size of the time frame in which the transcript data is used for analysis.
S: sliding step.
n: number of sliding step per window size, assumed to be whole number.
T: the length of one discussion session.
t: time since the beginning of the design discussion.

In PPT, the first step is to divide the discussion session into several intervals. If the interval size is too small, then there may be too few word occurrences to detect. If the interval is too big, then the granularity may be too coarse for analysis to be meaningful.

In order to address this, an improved approach is proposed which uses a "sliding window" to sharpen the granularity of the design process. We keep the windows sufficiently large to collect enough key terms but slide the window less than the length of one window size so that we get data at more time points.

Sliding window techniques are widely used for data mining and information retrieval [38, 39]. In our research, sliding windows were used to dynamically retrieve a team's preferences at a finer granularity without drawing more perturbations that occur when there is little design team discussion.

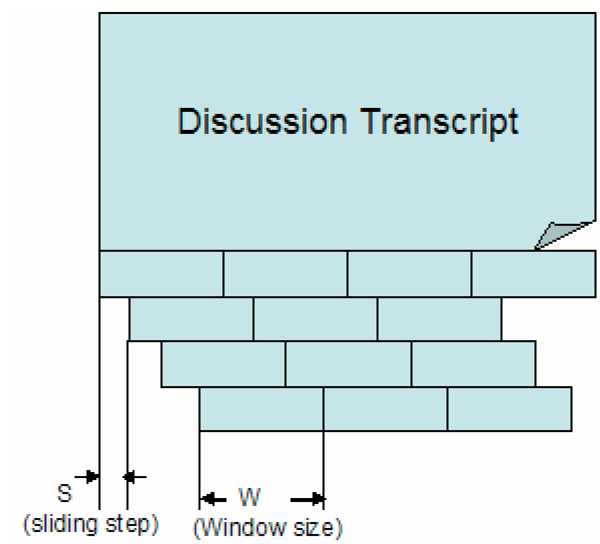


Figure 1. SPPT: Sliding Window for PPT

Figure 1 shows how sliding windows are used to extract the preference information from the transcript of the design team discussion.

The basic steps for deriving SPPT include:

1. Choose a window size *W* which is appropriate as the interval size for PPT (e.g. 10 minutes) and a slider size *S* for granularity requirement (e.g. 2.5 minutes)
2. Start from the beginning of the transcript
3. Apply PPT with interval = W selected in step 1
4. Obtain preferential values for t=0, t=W, t=2*W, ….
5. Slide the starting point to t=S and use this point as the beginning of the transcript, and repeat steps 3 and 4

6. Repeat 5 until the slider reaches the end of the first interval t=W.
7. Iterate from Steps 2 to 6 until we obtain the preferential probabilities at this different time interval. The iteration process for SPPT is the same as PPT, and parameters are also guaranteed to converge.

In previous research, equal preferential probabilities were assigned when t=0 if we do not have any prior knowledge. In SPPT, there are several starting points, which are t=0, t=S, t=2*S, …, t=(n-1)*S.

For t=0, we can use the same preferential probabilities, but for the remaining starting points (t=S, t=2*S, …, t=(n-1)*S), we must obtain values through interpolation between the initial preferential probabilities and the preferential probability at t=W. This preferential probability at t=W is obtained by updating preferential probability using traditional PPT and assumption of preferential probability at t=0.

When employing PPT, either word utterances or the appraisal values can be used. In this paper, we are mainly comparing PPT with smaller interval and SPPT with bigger interval, in order to minimize the noises generated by other aspects (such as word utterances increase even when the designers utter some alternatives in a negative way), we use appraisal values for both PPT and SPPT.

Baseline: Preferential Probabilities extracted from design team Surveys (PPS)

A simple approach to assessing SPPT and PPT is to look for qualitative consistency between the original transcript and the preferential probabilities extracted from it [6]. "Qualitative consistency" means that what is said by the team members is generally reflected by the results of PPT, although this is, by definition, an informal measure. In early tests of PPT, results were largely consistent. However, more quantitative approaches for evaluation are important for two reasons. First, both the extracted preferential probabilities and qualitative readings are drawn from a single source – the discussion transcript. Second, a qualitative assessment cannot describe to what degree the preferential probabilities extracted with PPT correlate with the team's intentions.

This paper turns to surveys as a baseline source for assessing SPPT and PPT because they address both of the issues identified above, and because they are a common tool for eliciting preferences.

In the study, periodic surveys of individual and team ratings and rankings were conducted to elicit a design team's preferences in order to establish a baseline for comparison with SPPT and PPT. However, raw survey ratings and rankings are not in a form that is readily comparable to SPPT or PPT. This study uses a method called Preferential Probabilities extracted from Surveys (PPS) to translate ratings and rankings into preferential probabilities. This approach is drawn from the principle of maximum entropy [40, 41] so that surveyed preferential probabilities can be compared with those found using PPT. The principle of maximum entropy is chosen because it provides the least biased distribution for the given information. This method does not assume a distribution *a priori*. The distribution and parameters are calculated while maximizing information entropy so that it does not have any unknown parameters. The approach also considers the boundary constraint while applying the principle of maximum entropy, which generates distinctive distributions for different stated ratings.

Details of the derivation of PPS are given in [8] and are briefly summarized here. PPS assumes preference ratings can be random for both the individuals providing the rankings and the team overall, and applies the principle of maximum entropy to both the individual survey ratings and the team's ratings. Through simulation, statistical results are collected for estimating the preferential probabilities.

PPS includes three main steps:

1. Construct a probability distribution for each individual's rating preference for each alternative
2. Construct a probability distribution for the team's rating preference for each alternative
3. Generate team preferential probabilities through simulation

CASE STUDY

Case Description

The design team in the experiment was a group of three Industrial and Systems Engineering graduate students at the University of Southern California. One had a background in Mechanical Engineering with 7 years of work experience, one had a background in Electrical Engineering with 2 years of work experience, and the third had a background in Mechanical Engineering with no work experience. They had known each other for about 2 years but had never worked together on a design project before.

Before the experiment, each team member was given a think-aloud training exercise to practice naming each alternative with proper names rather than ambiguous pronouns ("this" or "that") in order to facilitate the tracking of design alternatives in the transcript. For example, a "glass carafe" can only be called "glass carafe", "glass pot", "glass coffee carafe", "glass coffee pot", "carafe A", "pot A", or "glass alternative." During the experiment, they discussed their preferences and rationale with each other until a consensus was reached. This discussion was recorded and transcribed. During the same exercise, they were asked to fill out surveys every ~10 minutes expressing their preference ratings for design choices. The experiment lasted 50 minutes, including 10 minutes for instruction and training, and 8 minutes for filling out 5 surveys during the session.

The team's task was to decide on two component selection design problems (a carafe and filter for a coffeemaker), each of which had three candidate alternatives that were provided a priori. Total cost for these two components together could not exceed $35:

Imagine you are a retired person who is a coffee connoisseur. Your day cannot begin until you make coffee each morning for you and your spouse. You are in good health but are not as strong or mobile as you were when you were younger. As a connoisseur, you prefer fresh ground coffee to instant coffee like Folger's, and you are well informed about the various types of gourmet coffee available, as well as the tools and equipment to prepare it. However, you are now on a fixed income and are conscious about how you spend your money which is why you make coffee at home rather than visit Peet's every morning.

One of the issues in studying preferences is that they may be ambiguous. Hey, Kulok and Lewis [42, 43] note that human designers may not be consistent when they state their preferences explicitly. This has the potential to make quantitative analysis of surveyed preferences difficult. The approach taken in this paper is to examine overall trends in preferences across a number of design alternatives, rather than assume that the findings for one alternative at one point in time are correct. In this study, only the carafe selection problem was taken as a case study. Table 1 lists the three carafe alternatives (glass, stainless-steel, and plastic). The designers were provided additional features and specifications that might play a role in their preferences for the carafe and filter in Table 1.

Studies of team discussion suggest that team members enter into discussion armed with only partial, independent knowledge of a topic, and group discussion can play a role in eliciting this partial knowledge so that better decisions may be made [44]. In order to encourage discussion among the participants and to better simulate a real-world team experience, information about the design choices was provided in the following ways:

Individuals were provided detailed information regarding only one of the three alternatives in order to simulate a realistic, partial knowledge scenario. Team members would then discuss product features in order to uncover information about the other alternatives.

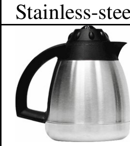
Surveys were conducted individually, and participants had no knowledge of their teammates' responses.

Individuals were encouraged to give a brief rationale for their rating and ranking to decrease the possibility of arbitrary ratings.

Both PPT and SPPT were applied to the resulting transcript. A base time interval of 10 minutes was used to calculate the preferential probabilities for two reasons. First, team members were asked to complete surveys every 10 minutes in the experiment, simplifying comparisons between extraction and survey results (PPS) [7, 8]. Second, earlier work showed that 10 minute intervals resulted in neither too many nor too few occurrences of terms.

In this application of PPT, the time interval was decreased to ¼ of the discussion session. Both SPPT and PPT results were compared with PPS. Though it would be possible to apply both SPPT and PPT only to word utterances, in this case appraisal values were also used in order to minimize the noise from other factors besides intervals.

Table 1. Design Information for Carafe Selection

Name/ID	Glass carafe	Stainless-steel carafe	Plastic carafe
Photo			
Description	Glass with warming plate	thermal-insulated stainless-steel	thermal-insulated plastics (inside glass)
Cost	$10.00	$20.00	$15.00
Warming plate cost	$5.00	0	0
Footprint size	Big	Small	Small
Fragility	Fragile	Strong	Fragile material inside
Durability (reliability)	Durable	Durable	Less durable
Heat retention	Good with heating plate	OK with double layers of steel	Good with mirror glass inside
Weight	Light	Heavy	Light
Portability	Not portable	Portable	Portable
Easy to clean	Easy to clean	Not easy to clean	Not easy to clean
Style and aesthetic value	Moderate attractive	Very attractive	Not attractive
Capacity	Can be designed as wanted Available for 2 cups and 6 cups	Can be designed as wanted Available for 2 cups and 6 cups	Can be designed as wanted Available for 2 cups and 6 cups
Spout	Not dribbles after pouring	Dribbles after pouring	Dribbles after pouring
Can tell how much coffee is left	Yes	No	No

Survey Ratings

The three designers are coded as X, Y and Z. Table 2 details designers' normalized survey ratings. Session 0 is the time when the design process had not started yet.

Table 2. Survey ratings from team members

Session	Designer	Carafe		
		Glass	Steel	Plastic
0	Designer X	0.5	0.3	0.2
	Designer Y	0.4	0.5	0.1
	Designer Z	0.4	0.5	0.1
1	Designer X	0.5	0.3	0.2
	Designer Y	0.5	0.4	0.1
	Designer Z	0.6	0.4	0
2	Designer X	0.5	0.2	0.3
	Designer Y	0.5	0.4	0.1
	Designer Z	0.6	0.4	0
3	Designer X	0.5	0.3	0.2
	Designer Y	0.6	0.3	0.1
	Designer Z	1	0	0

RESULTS

PPT with ¼ Session Interval on Appraisal

The carafe case study transcript was initially divided into 16 equal intervals (4 intervals in each 10 minute session). However, because there were no appraisals in the last session, only the appraisal values in the first 3 sessions are included.

In each interval, the segment of transcript was coded by following the steps prescribed in APA (Appraisal Preference Analysis) [7]. By identifying the appraisal clause in the transcript clause in the text [45], and by identifying how the linguistic appraisal reflects the preference values, we can obtain the qualitative appraisal coding of the preferences for the three alternatives, as shown in Table 3.

If the linguistic appraisal identifies a positive attitude toward the alternative or the attribute of the alternative, we increment the preference value for the increasing preference utility, For example, "glass+" means increased appraisal in the transcript for alternative "glass carafe", while "glass-" means decreased appraisal in the transcript for this alternative.

Following the steps in [7], negative appraisal values were converted to positive appraisal values which can be divided among other alternatives. The rationale behind this step is that a negative appraisal towards one alternative can be thought of as a positive appraisal towards all other alternatives. Without any other information, the principle of maximum entropy [40, 41] would assume that the shifting of preferential probabilities from an alternative with a negative appraisal to any other alternative should be equally likely. The converted frequencies of the appraisal values for each alternative in each interval are shown in Table 4.

Table 3. Appraisal Counts for Each Interval

Session	Carafe Alternatives					
	Glass		Steel		Plastic	
	+	-	+	-	+	-
0.25	0	0	0	0	0	0
0.5	1	0	0	0	0	0
0.75	2	1	1	1	1	0
1	3	0	6	6	3	5
1.25	0	0	1	0	0	2
1.5	0	0	0	0	0	0
1.75	6	0	4	0	0	0
2	0	0	0	0	0	0
2.25	0	0	0	0	0	0
2.5	0	0	0	0	0	0
2.75	6	1	1	0	0	0
3	0	1	0	2	0	0

Table 4. Converted Appraisal Frequencies

Session	Carafe Alternatives		
	Glass	Steel	Plastic
0.25	0	0	0
0.5	1	0	0
0.75	2.5	1.5	2
1	8.5	8.5	6
1.25	1	2	0
1.5	0	0	0
1.75	6	4	0
2	0	0	0
2.25	0	0	0
2.5	0	0	0
2.75	6	1.5	0.5
3	1	0.5	1.5

By applying PPT on the data in Table 4, the preferential probabilities calculated are shown in Table 6. The initial preferential probabilities at T=0 are 1/3 for all three alternatives.

SPPT with a larger interval window

Table 5 shows the appraisal values for three alternatives for a window size of one interval that is 10 minutes long, and slider that is 1/4 interval long (2.5 minutes).

When the window slides, the extracted preferential probabilities at T are based on the ones at T-1. Table 6 shows the preferential probabilities using SPPT.

Table 5. Converted Appraisal Count at "Sliding Window" intervals

Session	Carafe Alternatives		
	Glass	Steel	Plastic
1	12	10	8
1.25	13	12	8
1.5	12	12	8
1.75	15.5	14.5	6
2	7	6	0
2.25	6	4	0
2.5	6	4	0
2.75	6	1.5	0.5
3	7	2	2

DISCUSSIONS

Figures 2-4 show the preferential probabilities for each design alternative extracted using SPPT (with T/4 slider), PPT (with T/4 interval), and those translated using PPS from the survey. While the numeric values for SPPT, PPT, and PPS are not exactly the same, the trends between the three methods are generally similar. All three methods show the glass carafe as being the most preferred choice at the end of all three T intervals, the steel carafe as the second choice in all intervals, and the plastic carafe as the lowest ranked choice throughout the discussion. This is also consistent with a qualitative reading of the transcript.

Besides the general trend, we also notice big fluctuations on preference during Session 2 for both PPT and SPPT. A qualitative reading finds that the team was analyzing the features of steel carafe and plastic carafe in Session 2 after they had finished analyzing the glass carafe. The preference trends within Session 2 are also consistent with a qualitative reading of the transcript.

An interesting finding regarding PPT and SPPT for this case study is that PPT with smaller interval may overshoot when reflecting the preference fluctuations. For example, at t=1.25 T, PPT shows that steel carafe was the most preferred alternative and glass was the most preferred one at other time points. However, SPPT shows that in the whole process, the glass carafe was always the most preferred alternative. The interviews with the team members after the experiment confirmed that the most preferred alternative was always the glass carafe although sometimes in the discussion they thought steel carafe was also good alternative. Another interesting finding is that PPT with a smaller interval can reflect preference changes more quickly than SPPT. This is because PPT with smaller interval is based on the transcript in the last T/4 time interval, while SPPT is still based on the transcript in the last T time interval.

Table 6. Preferential Probabilities from SPPT, PPT, and PPS

Session	SPPT			PPT (1/4 T interval)			PPS		
	Glass	Steel	Plastic	Glass	Steel	Plastic	Glass	Steel	Plastic
0	0.333	0.333	0.333	0.333	0.333	0.333	0.496	0.479	0.025
0.25	0.395	0.321	0.284	0.333	0.333	0.333	-	-	-
0.5	0.457	0.308	0.235	0.424	0.288	0.288	-	-	-
0.75	0.518	0.296	0.186	0.411	0.266	0.323	-	-	-
1	0.580	0.284	0.136	0.431	0.411	0.159	0.659	0.322	0.019
1.25	0.550	0.368	0.082	0.325	0.475	0.200	-	-	-
1.5	0.475	0.431	0.094	0.325	0.475	0.200	-	-	-
1.75	0.612	0.374	0.014	0.633	0.308	0.059	-	-	-
2	0.610	0.356	0.034	0.633	0.308	0.059	0.679	0.271	0.050
2.25	0.658	0.291	0.051	0.633	0.308	0.059	-	-	-
2.5	0.638	0.309	0.053	0.633	0.308	0.059	-	-	-
2.75	0.793	0.140	0.068	0.796	0.126	0.077	-	-	-
3	0.800	0.116	0.085	0.382	0.252	0.366	0.886	0.098	0.016

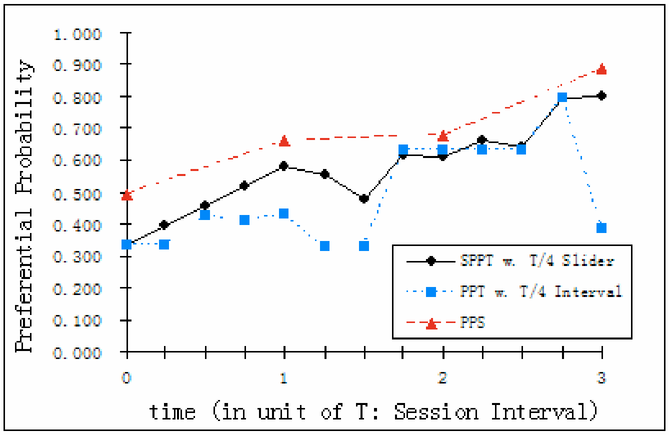


Figure 2: Time variation of preferences for the Glass Carafe

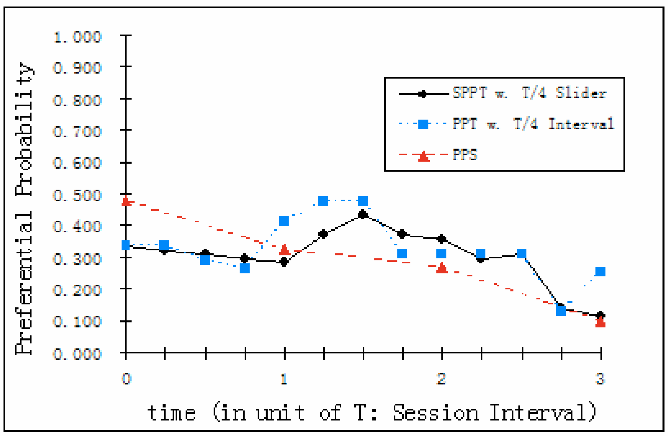


Figure 3: Time variation of preferences for the Steel Carafe

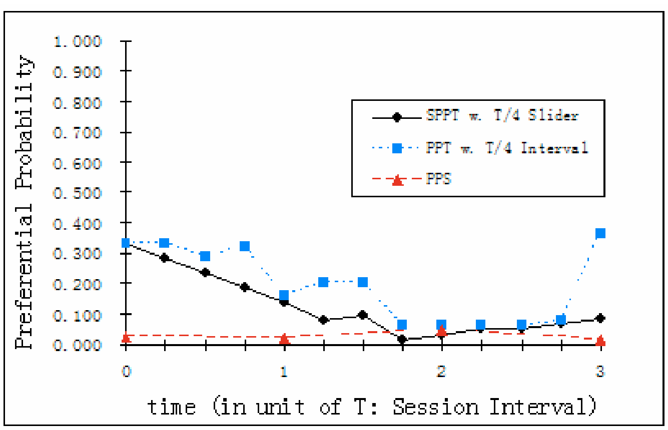


Figure 4: Time variation of preferences for the Plastic Carafe

In this study, the Pearson product-moment correlation coefficient was calculated to assess whether the correlations between the approaches is statistically significant. The Pearson Correlation coefficient can range from +1 to -1. The closer the correlation coefficient is to the +1/-1, the more two variables are correlated positively/negatively.

Since preferential probabilities from surveys were only calculated at the end of each session, only the preferential probabilities from SPPT and PPT at the end of each session (t=1,2,3) were compared with PPS. The Pearson correlations between these three data sets (SPPT, PPT, PPS) are shown in Table 7. We can see SPPT has much better correlation with PPS than using smaller interval for PPT. However, all 3 correlations are statistically significant. When we use small intervals for PPT, it is possible that in some intervals, there are too few word utterances or word appraisals. These could cause noise when applying PPT and PPT values at some certain points have big errors (such as at Time=3T for PPT with T/4 Interval). Using a bigger window size but sliding in a smaller step, we can minimize the perturbations caused when there is little team discussion or few appraisals.

Table 7. Pearson correlations

	SPPT (T/4 slider)	PPT (T/4 interval)	PPS
SPPT (T/4 slider)	-	0.692 p=0.039	0.986 p=1e -6
PPT (T/4 interval)	-	-	0.693 p=0.039
PPS	-	-	-

CONCLUSION

This paper presents a sliding window approach to improve the granularity when extracting the probabilistic preference information from the transcripts of team discussion. It follows the basic idea of PPT and does not require explicit input from the designer about his or her preferences. The approach treats a discussion as a "bag of words" that contains all of the terms used by the designers, and does not take into consideration what individuals say or how individuals' comments should be aggregated. The preferences extracted are based on the discussion covered by the sliding window, while the granularity of preference changes is controlled by the sliding steps.

This approach provides a better time-based representation of preference-related information which can illustrate how design selection changes over time. It can help us understand the design process in a finer granularity. This work may contribute to understand the evolving nature of a team's preferences over the life of a project. This paper also established graphical and quantitative consistency between the preferential probabilities extracted from a transcript and those translated from surveys. The approaches proposed in this study as well as in previous work [6, 7] provides a quantitative and implicit way to understanding the qualitative design process in a low overhead extraction way. This study describes how the team's preferences on selection evolve throughout the design process. An interesting topic in future research would be studying the rationale a design team selects a specific alternative or a set of specific alternatives over another by also capturing attribute information.

This work investigated a small group discussing a simple task. Future work should include testing SPPT and PPT on a variety of other cases and scenarios that take into consideration a range of team sizes, backgrounds and design task complexity. Many design problems in practice involve larger teams addressing more complex problems. Future work may consider the role of group dynamics and team membership roles, such as authority, decision-making styles, and dominance of individual team members.Both SPPT and PPT can be applied to transcripts from any size discussion group, but there may be possible effects of group size on quantity and quality of discussion.

ACKNOWLEDGMENTS

The work described in this paper was supported in part by the National Science Foundation under Award CMMI-0900255. The opinions, findings, conclusions and recommendations expressed are those of the authors and do not necessarily reflect the views of the sponsors..

REFERENCES

1. Pugh, S., *Total design: integrated methods for successful product engineering*. 1991, Wokingham, England: Addison-Wesley.
2. Hauser, J.R. and D. Clausing, *The House of Quality.* Harvard Business Review, 1988. **66**(3): p. 63-73.
3. Hazelrigg, G.A., *A framework for decision-based engineering design.* Journal of mechanical design, 1998. **120**(4): p. 653-658.
4. Wood, K.L. and E.K. Antonsson, *Computations with Imprecise Parameters in Engineering Design: Background and Theory.* ASME Journal of Mechanisms, Transmissions, and Automation in Design, 1989. **111**(4): p. 616-625.
5. Otto, K.N. and E.K. Antonsson. *The Method of Imprecision Compared to Utility Theory for Design Selection Problems*. in *ASME 1993 Design Theory and Methodology Conference*. 1993.
6. Ji, H., M.C. Yang, and T. Honda, *A Probabilistic Approach for Extracting Design Preferences from Design Team Discussion.* Proceedings of ASME 2007 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference, 2007.
7. Honda, T., et al. *A Comparison of Formal Methods for Evaluating the Language of Preference in Engineering Design*. in *International Design Engineering Technical Conferences & Information in Engineering Conference*. 2010. Montreal, Canada.
8. Ji, H., *Extraction of Preferences from Early Stage Engineering Design Team Discussion*, in *Department of Industrial and Systems Engineering*. 2008, University of Southern California: Los Angeles, CA, USA.
9. von Neumann, J. and O. Morgenstern, *Theory of Games and Economic Behaviour*. 2nd Edition ed. 1947, Princeton: Princeton University Press.
10. Howard, R.A. and J.E. Matheson, *The Principles and Applications of Decision Analysis*. 1984, Menlo Park, CA: Strategic Decision Group.
11. Keeney, R.L. and H. Raiffa, *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*. 1976, New York: Wiley.
12. Thurston, D., *A Formal Method for Subjective Design Evaluation with Multiple Attributes.* Research in Engineering Design, 1991. **3**(2): p. 105-122.
13. Thurston, D. and A. Locasio. *Multi-attribute optimal design of structural dynamic systems*. in *the 1992 ASME Design Theory and Methodology Conference*. 1992. Scottsdale, AZ, USA.
14. Tribus, M., *Rational Descriptions, Decisions, and Designs*. 1969, New York: Pergamon Press.
15. Otto, K.N. and E.K. Antonsson, *Trade-Off Strategies in Engineering Design.* Research in Engineering Design, 1991. **3**(2): p. 87-104.
16. Krantz, D.H., et al., *Foundations of Measurement Volume 1*. Vol. I. 1971, New York, NY, US: Academic Preess.
17. Saaty, T.L., *Fundamentals of Decision Making and Priority Theory With the Analytic Hierarchy Process*. Analytic Hierarchy Process Series. Vol. 6. 2000, Pittsburgh: RWS Publications.
18. Wang, J., *A Fuzzy Outranking Method for Conceptual Design Evaluation.* International Journal of Production Research, 1997. **35**(4): p. 995-1010.
19. Scott, M.J. and E.K. Antonsson, *Aggregation functions for engineering design trade-offs.* Fuzzy Sets and Systems, 1998. **99**(3): p. 253-264.
20. Green, P.E. and V. Srinivasan, *Conjoint Analysis in Marketing: New Developments with Implications for Research and Practice.* Journal of Marketing, 1990. **54**(4): p. 3-19.
21. Hensher, D.A. and L.W. Johnson, *Applied Discrete Choice Modeling*. 1981, New York: Halsted Press.
22. Ben-Akiva, M. and S.R. Lerman, *Discrete Choice Analysis*. 1985, Cambridge, Massachusetts: The MIT Press.
23. Vroom, V. and A. Jago, *The New Leadership: Managing Participation in Organizations*. 1988, Upper Saddle River, NJ: Prentice-Hall.

24. Arrow, K.J. and H. Raynaud, *Social Choice and Multicriterion Decision-Making*. 1986, Cambridge, MA, US: The MIT Press.
25. Arrow, K.J., *Social Choice and Individual Values*. 2nd ed. 1970, New Haven, CT, US: Yale University Press. 126.
26. Dym, C.L., W.H. Wood, and M.J. Scott, *Rank ordering engineering designs: pairwise comparison charts and Borda counts.* Research in Engineering Design, 2002. **13**(4): p. 236-242.
27. Keeney, R.L., *A Group Preference: Axiomatization with Cardinal Utility.* Management Science, 1976. **23**(2): p. 140-145.
28. Jabeur, K., J.-M. Martel, and S.B. Khelifa, *A Distance-Based Collective Preorder Integrating the Relative Importance of the Group's Members.* Group Decision and Negotiation, 2004. **13**(4): p. 327-349.
29. See, T.-K. and K. Lewis, *A Formal Approach to Handling Conflicts in Multiattribute Group Decision Making*. Journal of Mechanical Design, 2006. **128**(4): p. 678-688.
30. Dong, A., *Concept Formation as Knowledge Accumulation: A Computational Linguistics Study.* Artificial Intelligence for Engineering Design, Analysis and Manufacturing, 2006. **20**(1): p. 35-53.
31. Yang, M.C., W.H. Wood, and M.R. Cutkosky, *Design Information Retrieval: A Thesauri-based Approach for Reuse of Informal Design Information.* Engineering with Computers, 2005. **21**(2): p. 177-192.
32. Ji, H., M.C. Yang, and T. Honda, *A Probabilistic Approach for Extracting Design Preferences from Design Team Discussion*, in *Proceedings of ASME 2007 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference*. 2007: Las Vegas, Nevada, USA.
33. Cross, N., H. Christiaans, and K. Dorst, *Analysing Design Activity*. 1996, Chichester: Wiley.
34. Halliday, M.A.K., *An introduction to functional grammar*. 3 ed, ed. C.M.I.M. Matthiessen. 2004, London: Arnold.
35. Martin, J.R. and P.R.R. White, *The Language of Evaluation: Appraisal in English*. 2005, New York: Palgrave Macmillan.
36. Dempster, A., N. Laird, and D. Rubin, *Maximum likelihood from incomplete data via the EM algorithm.* Journal of the Royal Statistical Society, Series B, 1977. **39**(1): p. 1-38.
37. Bilmes, J.A., *A Gentle Tutorial of the EM Algorithm and its Application to Parameter Estimation for Gaussian Mixture and Hidden Markov Models. Technical Report*. 1998, International Computer Science Institute: Berkeley, CA, USA.
38. Lee, C.-H., C.-R. Lin, and M.-S. Chen. *Sliding-window filtering: an efficient algorithm for incremental mining*. in *the tenth international conference on Information and knowledge management*. 2001. Atlanta, Georgia, USA.
39. Datar, M., et al., *Maintaining stream statistics over sliding windows.* SIAM Journal on Computing, 2002. **31**(6): p. 1794-1813.
40. Jaynes, E.T., *Prior Probabilities.* IEEE Transactions On Systems Science and Cybernetics, 1968. **sec-4**(3): p. 227-241.
41. Jaynes, E.T., *Information Theory and Statistical Mechanics.* Physical Review, 1957. **106**(4): p. 620-630.
42. Hey, J.D., *Do Rational People Make Mistakes?*, in *Game Theory, Experience, Rationality*, W. Leinfellner and E. Kohler, Editors. 1998, Kluwer Academic Publishers: Netherlands. p. 55-66.
43. Kulok, M. and K. Lewis, *Preference Consistency in Multiattribute Decision Making.* ASME Conference Proceedings, 2005. **2005**(4742Xa): p. 291-300.
44. Gigone, D. and R. Hastie, *The impact of information on small group choice.* Journal of personality and social psychology, 1997. **72**(1): p. 132-140.
45. Dong, A., M. Kleinsmann, and R. Valkenburg, *Affect-in-Cognition through the Language of Appraisals.* Design Studies, 2009. **30**(2): p. 138-153.