INFERENCE-ASSISTED CHOOSING BY ADVANTAGES

John Haymaker[1], Duen Horng Chau[2] and Bo Xie[3]

ABSTRACT

This paper presents an approach to leverage decision models and machine learning to assist designers facing a decision challenge to automatically recall relevant information from prior decisions. We specify a graph-based collaborative decision-making model based on the Choosing by Advantages (CBA) methodology that unifies elements of rationale involved in the decision making process (e.g., team members, objectives, alternatives, attributes, advantages, importance) and the relationships among them in a format suitable for machine inference methods. We adapt the Belief Propagation (BP) graph-based inference method to analyze existing corpus of decision rationale to inform collaborative decision-making. We illustrate the CBA-BP algorithm with an explanatory example, describe initial implementation, and outline future work in populating databases of decision models, implementing the algorithm and associated user interfaces, and validating the efficiency and effectiveness of the algorithm and interfaces through user-centered testing.

KEYWORDS

Choosing By Advantages, Set-based Design, Time Compression, Decision Models, Data Mining, Inference, Belief Propogation, Case-Based Reasoning.

INTRODUCTION

Formal decision process models are beginning to change the way building professionals make decisions. For example, industry and researchers are actively investigating the Design Quality Indicator (Whyte and McGann, 2002), Multi Attribute Collective Design Assessment, and Decision Integration (Haymaker et al 2010), SPeAR (McGregor & Roberts, 2003), and Wecision (Abraham et al, 2013) to help design teams gather, structure, and communicate rationale. Lean Construction has embraced the Choosing by Advantage system (Macomber et al, 2006), as a preferred way to formalize the decision process and eliminate waste in this critical step of delivering customer value (Koskela et al, 2002).

However, currently decision process models rely almost entirely on human input to construct and organize the rationale needed to model each decision. The difficulty users encounter when entering well-structured rationale, and the as yet not validated benefit of doing so, means these systems are rarely implemented in practice (Conklin and Yakemovic 1991; Moran and Carroll 1996; Ishino and Jin 2002). Decision

[1] School of Architecture and Building Construction, College of Architecture, Georgia Institute of Technology, Atlanta, GA; email: haymaker@gatech.edu

[2] School of Computational Science and Engineering, College of Architecture, Georgia Institute of Technology, Atlanta, GA; email: polo@gatech.edu

[3] School of Interactive Computing, College of Computing, Georgia Institute of Technology, Atlanta, GA; email: boxie@gatech.edu

makers need methodologies and tools that make them both more effective, and more efficient (Senescu et al, 2013).

Semi-automated rationale construction methods present a potentially significant opportunity to support designers in making better decisions more quickly, accurately, and with greater confidence. This paper presents an approach that leverages decision models and machine learning to help designers with a decision challenge to draw on rationale of their own and others' past decisions. Our contributions include a graph-based collaborative decision-making model, and machine learning to work on this model. We develop a unifying model to capture pieces of information involved in the Choosing by Advantages decision-making processes (e.g., team members, objectives, alternatives, analyses) and the relationships among them. We explore how to adapt machine learning, specifically a graph-based inference method called *Belief Propagation*, to analyze existing corpus of decision rationale to inform collaborative decision-making. We present an explanatory example, describe initial and planned implementations, and discuss methods to measure their impact on decision-making efficiency and effectiveness. We target important challenges related to buildings, although the methods are likely more general.

DECISION PROCESS MODELS

Many decision models are possible; each with strengths and weaknesses depending on the challenge (Hazelrigg, 1999). Figure 1A illustrates many of the types of information and relationships that a decision maker might consider in a decision. An analysis of these methods is beyond the scope of this paper. Choosing by Advantages (Suhr, 1999) is a methodology that has gained popularity in the Lean Construction community (Macomber et al, 2003). CBA is a methodology containing many methods for making decisions of increasing complexity. In the CBA tabular method for decisions not involving money, a decision maker makes a decision by defining the alternatives, the criteria, the attributes of each alternative for each criteria, computing the advantages of alternatives, and then weighing the importance of advantages. The alternative with the most cumulative importance is the preferred alternative. This paper explores the potential of inference support for decision-making specifically in the context of the tabular method of CBA for decisions not involving money (Suhr, 2008.).

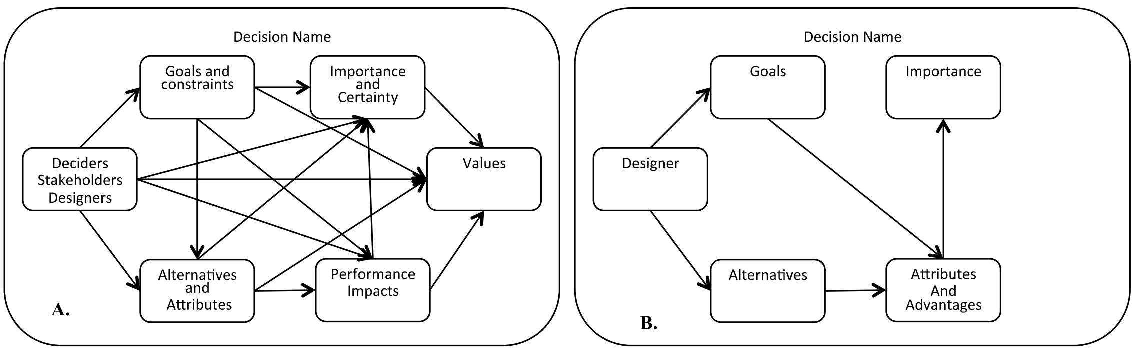


Figure 1A: A generic decision model illustrating many of the elements of rationale and relations relevant to a decision. B. The CBA Process addressed in this paper.

OPPORTUNITIES FOR INFERENCE ON CHOOSING BY ADVANTAGES DECISION MODEL

Figure 1 illustrates that decision models require the construction and relating of a great deal of rationale. Users of a decision rationale system could use support in many ways, including naming the decision, identifying appropriate stakeholders, designers, gatekeepers, and decision makers to include, in choosing and correctly defining the project objectives, in assigning priorities amongst these objectives, in generating alternatives and their attributes, in determining the advantages, and in assessing the importance of these advantages.

There are also different levels of complexity in the decisions AEC teams face. On one end of the spectrum are relatively simple decisions with few, well-defined goals, alternatives and performance impacts. For simpler decisions today's commercially available search and recommendation systems may be applicable. However, designers often face more complicated challenges, with a greater number of sometimes more qualitative goals, very large and continuous spaces of alternatives, complex and uncertain analyses, and conflicting priorities among stakeholders. This research investigates the applicability of Belief Propagation methods on these different levels of decision complexity.

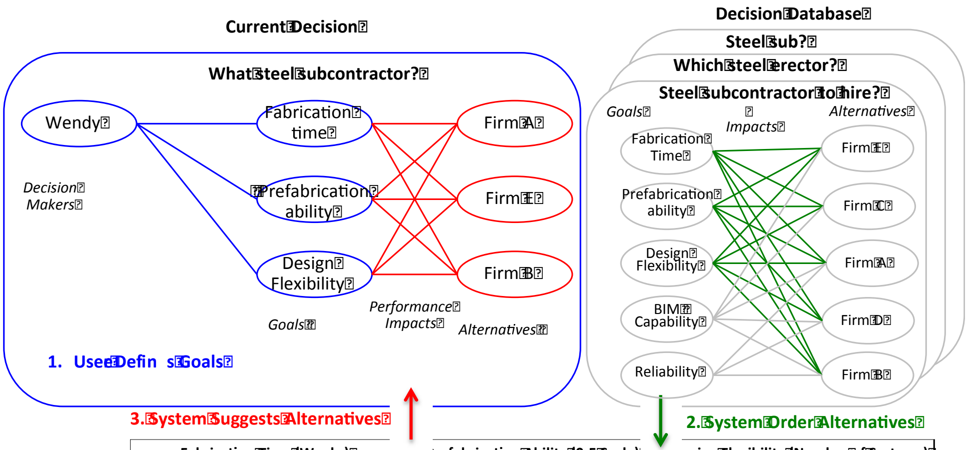
SCENARIO: WORKFLOW LEVERAGING MACHINE INFERENCE

We illustrate our envisioned workflow using the following scenario, where our system guides a decision maker, a student named Wendy, to quickly *identify a small set of favorable alternatives that are likely to maximize satisfaction of her goal preferences, and eventually choose the best alternative*. We will give the formal problem definition in the next section when we describe our methodology. Here, we use an example of choosing a trade subcontractor to explain how a decision maker and the system might interact, and illustrate some of the challenges and opportunities for integrating CBA and BP.

Step 1: Wendy needs to choose a subcontractor to erect steel on her project. In the beginning, she specifies several goals that she is aware of: the time it takes to fabricate an order, their ability to prefabricate large sections, and their flexibility to deal with unusual design conditions. She believes there may be more goals, but she decides to start with these and let the system help her refine these goals, and identify other relevant ones later.

Step 2: The system then (i) searches through its database of prior decisions to retrieve those that have similar decision names, e.g., containing or similar to the terms "steel" and "contractor"; (ii) and extract their alternatives that concern Wendy's goals.

Step 3: Our proposed machine inference approach, described in the next section, can infer which alternatives may best meet her needs. At the high level, the algorithm computes a "goodness" score for each alternative. Goodness is based on how strongly each alternative is associated with Wendy's goals. The three goals are the "sources of goodness", which diffuse to their neighbors, and those neighbors' neighbors iteratively.



	Fabrication Time (Weeks)			Prefabrication Ability (0-5 Scale)			Design Flexibility (Number of Systems)		
	Attribute (Weeks)	Advantage	Normalized Advantage	Attribute 0=bad, 5=good	Advantage	Normalized Advantage	Attribute (BIM Rank)	Advantage	Normalized Advantage
Firm A	4	8	1	5	3	1	83	19	.76
Firm E	8	4	.5	3	1	.33	89	25	1
Firm B	7	5	.625	5	3	1	64	-	0
Firm D	10	2	.25	3	1	.33	88	24	.96
Firm C	12	-	0	2	0	0	75	11	.44

Figure 2: In step one, the user defines and potentially weights goals. In step two, the system orders alternatives with respect to their performance on these goals. In step three, the system suggests the highest performing alternatives to the user.

One approach to determine the strength of association between a goal and an alternative is based on the alternative's advantage value. The algorithm can map such value into a number that governs how much "goodness" would spread through that edge. For example, consider the goal "Prefabrication Ability", one way is to normalize the five alternatives' advantage values, as shown in the table, to the (0, 1) range. This means that goodness from "Fabrication Time" would be spread to the five alternatives in these proportions.

In short, the algorithm determines each alternative's overall goodness, by aggregating the incoming goodness that it receives from the specified goals through multiple paths; the higher the aggregate goodness score, the better the alternative.

A major attractive aspect of our proposed graph-based inference approach is that it can handle user feedback. In the previous steps, Wendy's goal preferences and the desirability of alternatives' attributes can be easily mapped to edge weights in the graphical model to effect inference.

Our approach can also handle the deletion of alternatives. For example, Wendy realizes that the alternative "Firm A" is not feasible (Figure 3), due to unstated constraints. She deleted that alternative, which corresponds to removing its node, and its adjacent edges, from our graphical model. The inference re-runs to compute the goodness scores of the remaining two alternatives, which now have very close goodness scores. Which one is better?

Step 4: Our system helps Wendy by recommending other goals (for example, "BIM Capability") that other users have considered for these alternatives, but not yet considered by Wendy. Again, our system finds and ranks these potential goals in the database using the same method it uses for alternatives, as discussed in the previous steps. Wendy realizes that this "BIM Capabilities" goal is an important consideration. She accepts the recommendation and updates her weights on these four goals. Steps 5 and 6 repeat steps two and three to reorder and help the user select more alternatives, in this case "Firm D" which scores relatively well on fabrication time, prefabrication ability, and design flexibility, while also having good BIM capability.

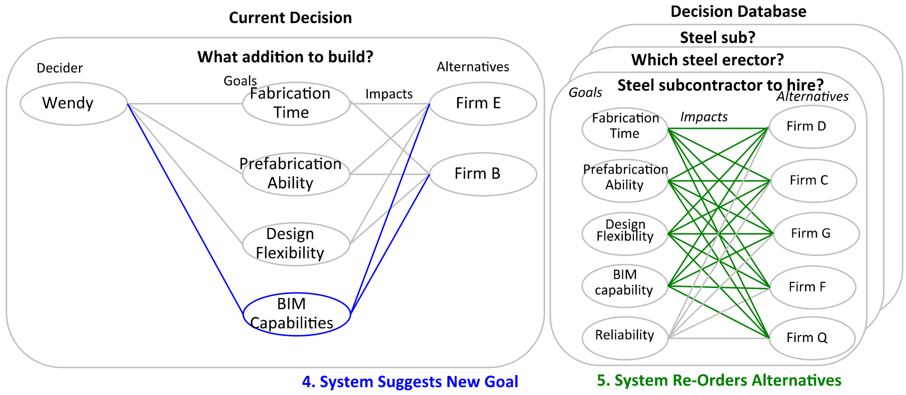


Figure 3: In step four the system suggests new goals. In step five, the system reorders alternatives.

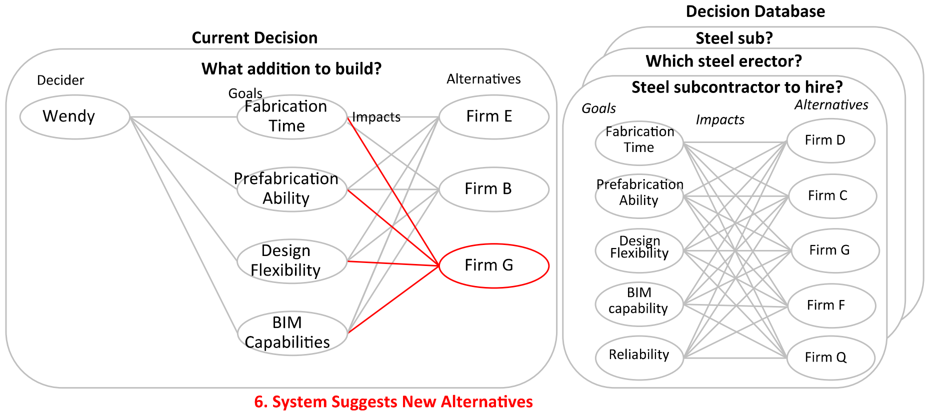


Figure 4: In step six, the system suggests new alternatives that map to the users revised, potentially reprioritized goals.

METHODOLOGY: FINDING RELEVANT ALTERNATIVES AND GOALS VIA GRAPH-BASED INFERENCE

The scenario above illustrates how we envision human and machine work together to locate good alternatives and identify relevant goals. Here we describe our proposed methodology in adapting the Belief Propagation machine learning method to help provide such inference support. We first give the problem definition, then we briefly describe the algorithm. Finally, we show a numerical example that demonstrates the results of applying the algorithm on a dataset from Wecision Enterprise developed in the scenarios described in Figures 2 - 4.

Problem Description. We can view the task of finding relevant alternatives as computing a **goodness** score (and its complementary badness score) for every alternative, given some goals that we are interested in (as in the scenario above). More technically, the problem is defined as:

Given: a undirected tri-partite graph of *stakeholders*, *goals* and *alternatives* (as shown in Figure 2), each viewed as a random variable $X \in \{x_g, x_b\}$, where x_g and x_b are the *good* and *bad* label respectively;

Goal: find the *goodness* for each alternative i, which equals the marginal probability of i being *good*, i.e., $P(X_i = x_g)$. An alternative's *goodness* and badness score sum to 1.

Computing the goodness for all alternatives is an NP-hard inference task (Yedidia et al, 2003). In practice, the Belief Propagation algorithm (BP) (Yedidia et al, 2003) is often used to approximately solve it and has been proven successful in many domains (e.g., malware detection (Chau et al, 2011), fraud detection (Pandit et al, 2007)). We adapt the algorithm here for decision making in building and construction. To the best of our knowledge, this has not been done before. We believe the algorithm has strong potential in supporting the decision making process, due to its scalability (time complexity linear to the number of edges in the graph), and its flexibility in incorporating user feedback as we illustrate in the above scenario (e.g., adding or removing alternatives or goals).

Algorithm Description. At the high level, the algorithm infers the label of a node from some prior knowledge about the node, and from the node's neighbors. This is done through iterative message passing between all pairs of nodes v_i and v_j. Let $m_{ij}(x_j)$ denote the message sent from i to j. Intuitively, this message represents i's opinion about j likelihood of being in class x_j. The prior knowledge about a node i, or the prior probabilities of the node being in each possible class are expressed through the node potential function $\phi(x_i)$. In our scenario, the random variables representing the three initial goals ("Fabrication time", "Prefabrication Ability", "Design Flexibility") have high prior values [4] which intuitively indicate high relevance or goodness (these goals are naturally relevant, since they were specified by the user. All other variables (nodes) have an unbiased prior (0.5).

At the end of the procedure, each alternative's goodness is determined. This goodness is an estimated marginal probability, and is also called belief, or formally $b_i(x_i)$ $(\approx P(x_i))$, which we can threshold into one of the binary classes. For example, using a threshold of 0.5, if the alternative belief falls below 0.5, the alternative is considered bad.

In details, messages are obtained as follows. Each edge e_{ij} is associated with messages $m_{ij}(x_j)$ and $m_{ji}(x_i)$ for each possible class. Each message vector m_{ij} is normalized over j, so that it sums to one. Normalization also prevents numerical underflow. Each outgoing message from a node i to a neighbor j is generated based on the incoming messages from the node's other neighbors. Mathematically, the message-update equation is:

$$m_{ij}(x_j) \leftarrow \sum_{x_i \in X} \phi(x_i)\, \psi_{ij}(x_i, x_j) \prod_{k \in N(i)\backslash j} m_{ki}(x_i)$$

[4] Nodes' prior values are often assigned by the user; they are the user's subjective belief of whether the nodes are relevant, before any computation or inference is performed. For example, a prior value close to 1 means the user strongly believes that node is relevant.

where $N(i)$ is the set of nodes neighboring node i, and $\psi_{ij}(x_i, x_j)$ is called the edge potential; intuitively, it is a function that transforms a node's incoming messages into the node's outgoing ones. Formally, $\psi_{ij}(x_i, x_j)$ equals the probability of a node i being in class x_i given that its neighbor j is in class x_j. In our case, since each edge has a weight (e.g., goal weight), this probability is scaled by that weight value.

The algorithm stops when the beliefs converge (within some threshold; 10^5 is commonly used), or a maximum number of iterations has finished. The node beliefs are then determined as follows, where k is a normalizing constant:

$$b_i(x_i) = k\phi(x_i) \prod_{x_j \in N(i)} m_{ji}(x_i)$$

Numerical Example. Here, we discuss initial efforts to apply and test the algorithm. Our method requires a collection of similar and appropriately structured decisions in a database, and there are currently a very limited number of decision types for which we have collected a large enough set of CBA models. We introduce here a scenario based on dataset from Wecision Enterprise developed in a class taught by the first author in which students were asked to construct their own decision models about how to mass a new addition onto a campus building. This gave us a collection of several models that attempt to answer a similar question with which to begin to explore the feasibility of the algorithm. As this was student work, we needed to adjust some models to improve their semantic and syntactic quality. We implemented the Belief Propagation algorithm in Java 1.6.

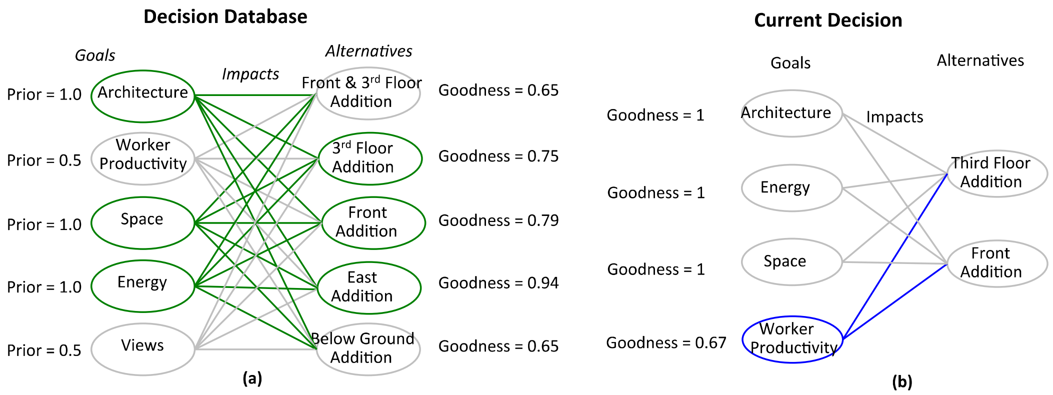


Figure 5. Example results from applying Belief Propagation on a class dataset. (a) Computes goodness scores for all alternatives and returns top ones as recommendations. (b) We get goal commendations "for free" since goodness scores are computed for all goals and alternatives every time the algorithm is run.

In this new scenario Wendy specified three goals ("Architecture", "Space", and "Energy"), and she wanted to find alternatives that maximize her goal preferences. The algorithm helps by computing the goodness for all alternatives available in the decision database, as shown in Figure 5a (similar to Figure 2, but annotated with algorithm inputs and outputs). The algorithm first sets the three goals' prior values to 1 (to indicate "high relevance"), and all other goals' to 0.5 (indicate "unknown relevance"), then uses the two equations described above to compute every alternative's goodness score. The three alternatives "3rd Floor Addition", "Front

Addition", and "East Addition" have the highest scores and are thus recommended to Wendy.

Later in this scenario (similar to Figure 3), Wendy wants the system to help identify more relevant goals. This means we are now interested in goals' goodness scores, instead of alternatives' scores. Since the algorithm always computes the goodness scores for all noes in the graph (i.e., all alternatives and all goals), we merely need to rank the goals by their scores to locate the next relevant goal (in this example: "Worker Productivity", as show in Figure 5b).

NEXT STEPS: EXTENDING AND VALIDATING THE MODEL

Our next step is to deploy our method and study how it may help decision makers in real-world tasks and decisions. We plan to conduct longitudinal studies to shadow decision makers during the course of projects that they work on, to obtain both quantitative and qualitative data and evidence, to help us better understand what kind of tasks our method may help with, to what extent; whether it is intuitive to use; are there drawbacks, etc. In the sections above, we focused on scenarios where we help the user find relevant alternatives and goals. However, we believe our method can extend to support locating other entity types as well, such as to find team members for a decision team. Below we describe our hypotheses and metrics for such evaluation to measure success. We envision conducting comparative studies where users will be divided into three groups (conditions), who will use different tools:

- **Condition C**: Conventional tools, such as spreadsheet
- **Condition W**: Wecision without Belief Propagation inference support
- **Condition WB**: Wecision with Belief Propagation inference

Our high-level hypothesis is that by leveraging a repository of *past* decision processes, we may help current decision makers make *better* decision. Table 1 describes metrics by which to measure improvements (i.e., what "better" means), and the more detailed, lower-level hypothesis that we have in mind (">" means greater, faster, more, etc.):

Table 1. Planned validation scheme for CBA-BP

Hypothesis	**Expected results**	**Quantitative or qualitative measure?**
Time to complete whole decision process	C > W > WB	Quantitative
Time to find desired number of goals	C > W > WB	Quantitative
Number of goals found	WB > W > C	Quantitative
Quality of goals defined (e.g., more details such as defining measurement units, using more words, pictures, URLs)	WB> W > C	Qualitative (may seek domain experts to help evaluate participants' data)
Time to find desired number of alternatives	C > W > WB	Quantitative
Number of alternatives found	WB > W > C	Quantitative
Quality of alternatives defined (e.g., more details such using more words, pictures, URLs, better advantages)	WB > W > C	Qualitative (may not need to rely on domain experts since main evaluation is performed in analysis step)
Time to assemble desired number of team members	C > W > WB	Quantitative
Qualification of team members found	WB > W > C	Qualitative (if decision makers can find the right team members, the team will be more collaborative and be able to produce better work)

CONCLUSION AND NEXT STEPS

We present a new approach leveraging decision models and machine learning to help guide users to leverage past decision rationale to support current decision-making. We describe the model of decisions necessary to support machine inference and the algorithm used to mine these decisions. We illustrate the use of this algorithm on a building massing decision scenario to explain how the algorithm works in the context of a decision. We adapted the Belief Propagation machine learning method, not to our knowledge applied to decision models, and demonstrated a large potential for providing assistance to decision makers.

Our next steps are to extend and integrate our prototype, to improve the corpus of decisions modeled therein, and to validate the impact that machine-learning assisted decision tools can have on decision methods. We will extend and integrate our prototype into the Wecision decision-making platform (DPI, 2013). Many types of algorithms, and approaches to providing support are possible. For example, one problem with any search algorithm is consistent and well structured information - one decision maker might call a goal “energy efficiency” while another might refer to a similar goal as “energy savings”. This problem could be addressed through requiring that more structured decision rationale be built by professionals, or by employing similarity algorithms that provide an additional layer of belief about the similarity of two text strings. Additional research is needed to determine the relative costs and benefits of different approaches. As long as the value of structuring decision rationale does not outweigh the immediate cost of recording the rationale, designers will not take the time to build good decisions.

We wish therefore to systematically test the ways that inference methods such as Belief Propagation can help improve the efficiency and effectiveness if decision making. Our hypothesis is that different types of data mining approaches could help designers decide what to call their decisions; the constraints they should define, goals they should consider, which goals should be most important, the team members they might include, the alternatives they could consider, the likely performance of these alternatives, and the ways different stakeholders are likely to react to potential outcomes of decisions. Different data mining algorithms will enable different types of decision makers to perform more or less effectively on different types of decision challenges. We propose a sophisticated evaluation methodology is needed to determine which methods provide the best guidance to which types of decision makers, on which types of challenges. These studies will consist of between-subject or within-subject user studies in which students and professional users will be asked to make decisions with and without various machine-learning algorithms. Decision efficiency will be evaluated in terms of the time and resources required to arrive at a decision, while effectiveness will measure issues related to the time required to complete the task, the number of people it involved, the number and range of alternatives explored, the certainty of the analyses performed, the value of the alternative selected and the satisfaction of the decision makers and stakeholders involved in the decision.

REFERENCES

Abraham, K., Lepech, M., and Haymaker, J. (2013). “Selection and Application of Decision Methods On a Sustainable Corporate Campus Project,” Proc. 21st

Annual Conference of the International Group for Lean Construction, IGLC, Fortaleza, Brazil.

Chau, D., Nachenberg, C., Willhelm, J., Wright, A., & Faloutsos, C. (2011). Polonium: Tera-scale graph mining and inference for malware detection. In Proccedings of SIAM International Conference on Data Mining (SDM) 2011.

Conklin, E. J., and Yakemovic, K. C. B. (1991). "A process-oriented approach to design rationale." Human Computer Interaction, 6, 357-391.

Design process Innovation (2013). Wecision. http://wecision.com [last accessed March 28, 2013].

Haymaker, J. R., Chachere, J. M., and Senescu, R. R. (2011). "Measuring and improving rationale clarity in a university office building design process." J. Arch. Eng., 17(3), 97-111.

Hazelrigg, G. (2003). "Validation Of Engineering Design Alternative Selection Methods," Eng. Opt., 2003, Vol. 35(2), pp. 103–120.

Ishino, Y., and Jin, Y. (2006). "An information value based approach to design procedure capture." Adv. Eng. Inf., 20(1), 89-107.

Koskela, L.; Howell, G.; Ballard, G.; Tommelein, I. (2002). "Foundations of Lean Construction". In Best, Rick; de Valence, Gerard. Design and Construction: Building in Value. Oxford, UK: Butterworth-Heinemann, Elsevier. ISBN 0750651490.

Macomber, H., Howell, G., and Barberio, J. (2007). "Target-Value Design: Nine Foundational and Six Advanced Practices for Delivering Surprising Client Value," White Paper, http://www.leanproject.com/wp-content/uploads/2011/10/3-Target-Value-Design-LPC.pdf [Accessed Jan 05, 2013].

McGregor, A. I., and Roberts, C. (2003). "Using the SPeARTM assessment tool in sustainable masterplanning." Presented at the US Green Building Conference – Greenbuild 2003, Pittsburgh, USA, 10 pp.

Moran, T. P., and Carroll, J. M. (1996). Design Rationale: Concepts, Techniques, and Use, Lawrence Erlbaum Assoc., Mahwah, NJ.

Pandit, S., Chau, D. H., Wang, S., & Faloutsos, C. (2007, May). Netprobe: a fast and scalable system for fraud detection in online auction networks. In Proceedings of the 16th international conference on World Wide Web (pp. 201-210). ACM.

Senescu, R., J. Haymaker, Meza, S., and M. Fischer, (2011). "Design Process Communication Methodology: Improving the Efficiency and Effectiveness of Collaboration, Sharing, and Understanding," accepted in ASCE Journal of Architectural Engineering.

Suhr, J. (1999). The Choosing By Advantages Decisionmaking System, Quarum Books, Westport, Ct.

Suhr, J. (2008). The Choosing By Advantages Tabular Methods and Money Decisions, Decision Innovations, Ogden Utah.

Whyte J.K., and Gann D., M (2003). "Design Quality Indicators: work in progress," Building Research & Information, 31(5), September–October, 387–398.

J. Yedidia, W. Freeman, and Y. Weiss, (2003). "Understanding belief propagation and its generalizations," Exploring artificial intelligence in the new millennium, vol. 8, pp. 236–239.