Qualitative Reasoning About Constraint Activity Using
Monotonic Influence Diagrams

Nestor F. Michelena
Post Doctoral Fellow

Alice M. Agogino
Professor

The Robotics Institute
Carnegie Mellon University
Pittsburgh, PA 15213

Department of Mechanical Engineering
University of California
Berkeley, CA 94720

Abstract

Monotonic influence diagrams (MID) are proposed for representing and manipulating qualitative and mathematical relationships between variables and constraints in order to design from physical principles. The theory of MID’s is based on a graph-theoretic representation of an optimization problem which can be topologically transformed as a means of solving the problem and exploring variable-objective-constraint relationships. Monotonic influence diagrams are a synthesis of influence diagrams and monotonicity analysis. Formally, a monotonic influence diagram is a directed graph consisting of nodes and arcs. The nodes represent design variables and the arcs reveal their relationships. Nodes in a MID can represent either deterministic or uncertain quantities. A deterministic qualitative relation between two variables is given by the sign of the partial derivative of the function defining one of the variables with respect to the other variable. A probabilistic qualitative relation is defined in terms of a constraint on the joint probability distribution of the variables. Only deterministic quantities and relationships will be addressed in this paper. Topological transformations such as arc reversal and node removal allow us to determine qualitative relations between constrained design variables and the objective function to be minimized or maximized. In this sense, MID’s provide a reasoning mechanism about constraint activity which entails explicit reasoning about inequality constraints, so candidates for active constraints or flaws in the problem formulation can be detected.

Introduction

Any attempt to develop a tool for engineering design entails having to manage design constraints, pursue design goals, consider uncertain engineering data and, most of the time, deal with an incomplete mathematical model. The last two aspects are more relevant during the early stages of the design process and are the most difficult to deal with. On the other hand, it is well known that the quality of the analysis at these early stages will minimize any effort involving redesign and reduce downstream lifecycle costs, including manufacturing and maintenance.
Qualitative probabilistic networks (QPN) [Wellman: 1990a&b] are abstractions of influence diagrams that encode constraints on the probabilistic relations among variables rather than precise numeric distributions. Qualitative relations express monotonicity constraints on direct probabilistic relations between variables, or on interactions among direct relations. Like influence diagrams, qualitative probabilistic networks facilitate graphical inference: qualitative relations of interest can be derive via graphical transformations of the network model.

Qualitative optimization, in the form of monotonicity analysis [Papalambros and Wilde: 1988], has been used to simplify the solution of, or completely solve in closed form, optimal design problems which would have otherwise required extensive numerical computation. Qualitative optimization has been used to decompose an optimization problem into a reasonable number of smaller subproblems which are readily solved. By analyzing the solutions of these smaller problems, the solution to the original problem can be identified. Principles of monotonic relationships between variables have also been used in AI research to reason qualitatively about physical systems based on first principle models [Bobrow: 1985].

In the firstPrince system, Cagan and Agogino [1991] have gone one step further in developing a design methodology based on qualitative optimization and symbolic algebra that innovates new optimal designs from an original prototype. Critical variables are identified and expanded to create new variables that promise to improve the design relative to the objective. The problem is then reformulated with the newly created variables and constraints are automatically modified or added as needed, subject to user-specified boundary conditions.

Monotonic influence diagrams have been proposed for both knowledge representation and qualitative and mathematical functional reasoning about the constraints and goals of an engineering optimal design problem [Michelena and Agogino: 1992a&b; Michelena: 1991]. A system represented by a monotonic influence diagram can contain both deterministic and random variables and relationships.

Formally, monotonic influence diagrams are directed graphs $G = (V, A)$ consisting of the sets of nodes $V$ and arcs $A$. In deterministic MID’s, we can differentiate two types of nodes, design variables $D$ and the objective node $f$, so $V = D \cup \{ f \}$. In general, arcs can represent deterministic relationships between nodes (such as mathematical functions) as well as constraints on the joint probability distribution of the variables, as in qualitative probabilistic networks [Wellman: 1990a&b].

Figure 1 shows the representation of a design problem with two design variables. The design objective is to minimize the weight of a circular beam, represented by the diamond-shaped node $w$. The design variables are the beam diameter and maximum stress, represented by circular nodes $d$ and $s$, respectively. A sign on an arc represents the monotonicity of the function defining a variable with respect to (w.r.t.) another variable, and the “+” (or “−”) superscripts on variables in a node indicates that the variable is bounded above (or below). For example, the “+” superscript on $s$ in figure 1 indicates that the maximum stress in the beam has an upper bound, such as from a simple inequality constraint $s \leq s_u$. Also, the “+” sign on the arc from $d$ to $w$ in figure 1 means that the beam weight is a monotonically increasing function of its diameter, while the “−” sign on the arc from $d$ to $s$ means that the stress is a monotonically decreasing function of the diameter. In general, the monotonicity of a variable $z = f(x)$ w.r.t. $x$, where $f$ is a differentiable function, is given by the sign of the partial derivative, i.e.,

$$
\delta_{x,z} = \text{Sign} (\frac{\partial f}{\partial x})
$$

This concept can be extended to a non-differentiable function by considering the region of definition of the generalized gradient.

As shown in the next section, a MID can be topologically transformed to reduce the number of variables and equality constraints, so at some point in this process only the objective and constrained design variables will be present in the diagram. Further analysis involves the application of monotonicity analysis to determine active inequality constraints, i.e., inequality constraints forced to strict equality at optimality. An active constraint implies setting a variable to either its upper or lower bound, hence reducing the number of degrees of freedom (DOF) of the optimal design problem. A branching approach, assuming constraints on the variables as strict inequalities, i.e., as inactive, is also used to remove constrained variables. The removal of nodes and use of monotonicity analysis can reveal the same sets of active constraints (optimal solution candidates) detected by the SYMON (SYmbolic MONotonicity analyzer) program [Choy and Agogino: 1986].

Monotonicity analysis is a technique used to qualitatively analyze the interaction between the constraints of an optimal design problem at optimality in order to reduce the dimensionality of the problem, detect flaws in the problem formulation and gain qualitative insights about directions for improving the design. In constrained optimization problems, the optimal solution is often at a boundary of the feasible domain, forcing one or more inequality constraints to be active. If this information is known before numerical optimization is performed, the...
dimensionality of the problem and thus the computation time can be reduced.

The foundation for monotonicity analysis is two well-defined rules based on logic and qualitative reasoning of a well-constrained optimization problem. The rules of monotonicity analysis are a qualitative form of the Karush–Kuhn–Tucker optimality conditions in nonlinear programming [Karush:1939; Kuhn and Tucker: 1951]. A conceptual description of the first rule of monotonicity analysis states that if the objective is monotonic w.r.t. a variable, then there exists at least one active constraint which bounds the variable in the direction opposite of the objective. In the case of the example of figure 1, this would entail designing for maximum allowable stress, i.e., \( s = s_u \), as shown in figure 2(a). The second rule implies that if a variable is not contained in the objective function then it must be either bounded from above and below by active constraints or not actively bounded at all, the latter requiring that any constraint monotonic w.r.t. that variable must be irrelevant. Only the first rule is used with MID's since topological transformations such as those presented in the next section allow us to reduce the diagram to one with constrained variable nodes without predecessors and with the objective node as the unique successor. The terminal diagram for the circular beam design example is depicted in figure 2(b). This simple example has been presented only to illustrate the concepts involved in monotonic influence diagrams; however, MID's can be used to represent, analyze and solve more complex optimal design problems.

![Figure 2](image)

**Figure 2** (a) First Monotonicity Rule, (b) Reduced MID

We refer the reader to [Michelena and Agogino: 1992a] for details about the representation of optimal design problems by means of a monotonic influence diagram.

### Topological Transformations of Deterministic Monotonic Influence Diagrams

The reduction in the number of design variables and the transformation of a deterministic MID to a form suitable for monotonicity analysis is accomplished by repeated arc reversals and node removals similar to those used for evaluating probabilistic influence diagrams [Rege and Agogino: 1988; Shachter: 1986]. These transformations are based on the Chain Rule for Derivatives and the Implicit Function Theorem and provide the foundation for an algorithm for qualitative analysis and solution of optimization problems. The proofs for the topological transformations are presented in [Michelena and Agogino: 1992a].

#### Consistency of Monotonic Influence Diagrams

A MID \( G' = (V', A') \) is said to be consistent with another MID \( G = (V, A) \), where \( V' \) is a proper subset of \( V \), if and only if \( G' \) represents the problem \( \phi \), implicit in \( G \), after back-substituting the variables in the set \( \mathcal{V} \setminus V' \) using \( \mathcal{V} \setminus V' \) equality constraints in \( \phi \).

### Arc Reversal

Two cases can be differentiated for the transformation of arc reversal, depending on whether the tail node \(^1\) of the arc to be reversed has direct predecessors or not.

**Case 1:** Tail node has direct predecessors \(- P(x_i) \neq \emptyset \)

Consider the MID \( G = (V, A) \) in which \( x_i, x_j \in V, (x_i, x_j) \in A, x_i \) is not a multipath predecessor of \( x_j, P(x_i) \neq \emptyset \), i.e., the set \( P(x_i) \) of direct predecessors of node \( x_i \) is not empty, and \( \delta_{x_i x_j} \neq 0 \) (see figure 3). Then the MID \( G' = (V, A') \) is consistent with \( G \), where:

\[
A' = (A \cup \{(x_j, x_i)\}) \setminus \{(x_i, x_j)\} \cup \{(x_i, x_k) : x_k \in P(x_i)\}
\]

\[
A'' = \{(x_k, x_i) : x_k \in P(x_i) \setminus \{x_i\} \} \cup \{(x_k, x_i) : x_k \in P(x_i)\}
\]

That is, arcs are added from nodes in \( P(x_i) \setminus \{x_i\} \) to \( x_i \) and from nodes in \( P(x_i) \) to \( x_i \). The arc \( (x_i, x_j) \) is reversed and the arcs incident with \( x_i \) in the original diagram removed. The monotonicity for each arc is given by \( \delta_1 \). If \( x_k \in P(x_i) \cap P(x_j) \), then the arc from \( x_k \) to \( x_j \) has monotonicity \( (\delta_1 \otimes \delta_2) \oplus \delta_3 \), where \( \otimes \) is sign multiplication and \( \oplus \) is sign addition. These operations are defined in table 1.

![Figure 3](image)

**Figure 3** Arc Reversal when \( P(x_i) \neq \emptyset \)

\(^1\) The tail node of an arc from \( x \) to \( y \) is node \( x \)

\(^2\) \( \delta_2 \), for instance, accounts for the monotonicity of arc \( \langle x_k, x_i \rangle \) for some \( x_k \in P(x_i) \)
Table 1 \( \otimes \) and \( \oplus \) Operators for Combining Monotonicities

<table>
<thead>
<tr>
<th>( \otimes )</th>
<th>( \oplus )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( + )</td>
<td>( + )</td>
</tr>
<tr>
<td>( - )</td>
<td>( + )</td>
</tr>
<tr>
<td>( 0 )</td>
<td>( + )</td>
</tr>
<tr>
<td>( - )</td>
<td>( - )</td>
</tr>
<tr>
<td>( ? )</td>
<td>( + )</td>
</tr>
</tbody>
</table>

**Case 2: Tail node has no direct predecessors**  \( P(x_i) = \emptyset \)

Consider the MID \( G = (V, A) \) in which \( x_i, x_j \in V \), \( \{x_i, x_j\} \in A \), \( x_i \) is not a multipath predecessor of \( x_j \), \( P(x_j) = \emptyset \), and \( \delta_{x_i x_j} \neq 0 \) (see figure 4). Then the MID \( G' = (V, A') \) is consistent with \( G \), where:

\[
A' = (A \cup \{(x_i, x_j)\}) \setminus \{(x_i, x_j)\} \cup \{(x_i, x_j) : x_i \in P(x_j) \setminus \{x_i\}\}
\]

When no arc is incident with node \( x_i \), arcs are added from nodes in \( P(x_j) \setminus \{x_i\} \) to \( x_i \). The arc \( (x_i, x_j) \) is reversed and the arcs incident with \( x_j \) in the original diagram removed.

At the qualitative level only the monotonicities \( \delta_i \) are used for the topological transformations. On the other hand, at the functional level, the expressions for the partial derivatives can be used with symbolic algebra to modify the MID.

![Figure 4 Arc Reversal when \( P(x_i) = \emptyset \)](image)

**Node Removal**

Consider the MID \( G = (V, A) \) in which \( x_i \in V \), and \( S(x_i) \) and \( P(x_i) \) are the sets of direct successors and predecessors of \( x_i \), respectively (see figure 5). Then the MID \( G' = (V', A') \) is consistent with \( G \), where:

\[
V' = V \setminus \{x_i\}
\]

\[
A' = (A \cup \{A''\}) \setminus \{A''\}
\]

\[
A'' = \{(x_i, x_j) : x_i \in P(x_i), x_j \in S(x_i)\}
\]

\[
A''' = \{(x_i, x_j) : x_i \in P(x_i)\} \cup \{(x_i, x_j) : x_i \in P(x_i), x_j \in S(x_i)\}
\]

That is, on removal of node \( x_i \) the nodes in \( P(x_i) \) become the direct predecessors of nodes in \( S(x_i) \). If \( S(x_i) = \emptyset \), then \( x_i \) is barren and can be eliminated from the diagram.

After establishing the calculus of monotonic influence diagrams, we can now use those topological transformations to manipulate the original diagram. We want to modify the diagram to find monotonic dependencies between constrained variables and the objective function in such a way that some or all of the constrained variables can be set to their upper or lower limits. In order to automate the procedure we need an algorithm which will perform the necessary transformations given a monotonic influence diagram for the problem.

We have presented two algorithms in [Michelena and Agogino: 1992a]. Algorithm 1 performs transformations in order to remove nodes whose variables are, or are assumed to be, unconstrained. Algorithm 2 removes nodes whose variables need to be set at their limits and analyzes the problem for different combinations of inactive constraints. These algorithms do not consider the possibility of loss of information regarding constraint activity and monotonicities. In that regard, they can be modified according to the guidelines given by Michelena [1991].

**Loss of Qualitative Information for Deterministic Monotonic Influence Diagrams**

Conclusions drawn from monotonic influence diagrams (MID's) containing (qualitative) monotonic information are sensitive to the sequence of topological transformations followed to reveal them. Although diagrams obtained through a series of transformations are consistent, the value of the results may vary. For instance, two different sequences of transformations might result in different monotonicities for a given arc in the final diagrams. Results concerning constraint activity could also be different under different sequences of transformations even when the resultant diagrams contain the same nodes.
Similar problems manifest themselves in other knowledge representations that take into account some kind of qualitative information. Qualitative probabilistic networks [Wellman: 1990b], which consider constraints on the joint probability distributions, show this unfortunate behavior. Interval influence diagrams [Fertig and Breese: 1990], used to perform probabilistic reasoning using interval rather than point value probabilities, are subjected to degradation of bounds on the probability distributions. In both cases, successive transformations result in loss of independencies present in the initial diagram. No procedure has been proposed for minimizing qualitative ambiguity. An optimal sequence of transformations would depend not only on the topology of the representation but also on the qualitative influences between nodes.

For instance, let the diagram of figure 6(a) be a QPN where nodes x and y have both positive arcs to z and are marginally independent. Reversing the arc from y to z twice yields the diagram of figure 6(b). That is, only the positive qualitative influence from y to z remains unaltered. The resultant QPN captures neither the positive influence from x and y to z nor the marginal independence of x and y shown in figure 6(a). Now, let the same diagram (figure 6(a)) represent a MID where z is deterministically defined and functionally dependent on x and y. The same sequence of two consecutive reversals would produce the same diagram, hence preserving all initial relationships.

Thus, we can see that under two sequences of transformations, no-arc-reversal and two arc reversals, a MID yields stronger results than the equivalent QPN. This dissimilarity between MID’s and QPN’s could have been predicted by just observing the stronger nature of functional relationships with respect to probabilistic ones. However, MID’s containing only information about monotonicities, unlike numerical influence diagrams or MID’s that account for the mathematical functional form of relations between variables, still present the problem of loss of information under different sequences of transformations. For instance, suppose z is monotonically increasing w.r.t. y and x, and y is functionally increasing w.r.t. x. A MID representation is shown in figure 7(a). Reversing the arc from y to z twice yields the diagram of figure 7(b). The positive monotonicity between x and y has been lost at this qualitative level of analysis. In this case, a no-transformation operation is preferred to a two-arc-reversal operation, since the latter introduces ambiguity in the relation between x and y.

The issue of loss of information regarding constraint activity is also of paramount concern. Obviously, the appearance of undefined monotonicities (ambiguities) after a topological transformation is a primary cause for this mishap. However, a more subtle reason for loss of information concerning constraint activity can be found in how nodes are removed, regardless of the presence of undefined monotonicities.

The problem of loss of information has been approached from a constraint activity point of view. Given the task of removing a node by means of arc reversal(s) and node removal, we will select the sequence of transformations that give the stronger results concerning constraint activity. If two sequences are equivalent in this sense, the sequence that minimizes ambiguity has priority. We have derived conditions on both the topology of the diagram and the monotonicities between nodes connected to the node to be removed which assure that a statement about the activity of a constraint is still valid after a transformation is performed [Michelen: 1991].

![Figure 6 Loss of Information in a QPN After Two Arc Reversals](a) ![Figure 6 Loss of Information in a QPN After Two Arc Reversals](b)

![Figure 7 Loss of Information in a MID After Two Arc Reversals](a) ![Figure 7 Loss of Information in a MID After Two Arc Reversals](b)

Conclusions

Monotonic influence diagrams are an effective way of graphically representing optimal design problems and explicitly reasoning about physical principles captured as equality and inequality constraints. Transformations of the diagram can reveal features of the solution hidden in the original formulation, such as degeneracies or global monotonicities. This allows the designer to reformulate the problem, reducing its complexity because of fewer degrees of freedom. Qualitative reasoning about the interrelationships between any pair of design variables and/or performance characteristics is also possible.
MID's can be utilized both at the qualitative level, using only qualitative information such as the signs of the partial derivatives, and at the functional level, when the mathematical form of the equations for the derivatives is considered. Functional information reduces the amount of information concerning monotonicities lost at the qualitative level. Also, the topological transformations can easily be mapped to functional manipulations using symbolic algebra programs. The results obtained at the qualitative and functional levels are equivalent to those from the SYMON and SYMFUNE [Agogino and Almgren: 1987] programs, respectively. However, monotonic influence diagrams allow one to carry out both levels of analysis simultaneously and, therefore, are computationally more efficient.

Acknowledgements

This research was partially supported by the National Science Foundation PYI grant #8451622, a grant from Rockwell International and an equipment gift from the Digital Equipment Corporation. This support is gratefully acknowledged.

References


