Monotone and Almost-Monotone Boolean Functions: Frequent Maxmimally Infrequent Sets

Elke Eisenschmidt and **Utz-Uwe Haus**
Institute for Mathematical Optimization
Department of Mathematics
Otto-von-Guericke-Universität Magdeburg
Universitätsplatz 2, 39106 Magdeburg, Germany

Enrico Schmidt and **Julia Sämann**
Regulatory Networks
Center for Biological Systems Analysis (ZBSA)
Albert-Ludwigs University Freiburg
Habsburgerstrasse 49, 79104 Freiburg, Germany

1 Introduction

Systems Biology is quickly becoming a major focus of applied and computational mathematics research. While traditionally numerics and continuous modeling techniques have been used, in recent years it has become apparent that qualitative modeling, data quantization and data mining questions play a central role for the field.

Monotone Boolean functions are ubiquitous in these models, and, surprisingly, often both the function and its dual simultaneously convey practically useful information: Consider the hypergraph defined by the metabolites interacting in a metabolic network, expressed as the stoichiometric matrix. Then the so-called *elementary modes* are nothing but the extreme rays of the rational polyhedral cone given by the matrix, and their support patterns form a hypergraph. Every transversal of this hypergraph is a so called *cut set*, and the set of all (minimal) cut sets is of central interest when analyzing failure modes of the system under consideration (Haus, Klamt, & Stephen 2008). Many other biological phenomena can also be modeled using hypergraphs, or would profit from the combinatorial richness that simple graph approximations cannot provide (Klamt, Haus, & Theis 2009).

Monotone functions also arise from non-hypergraph models: In the context of ancestral genome reconstruction one question is whether a matrix whose columns are genomic markers and whose rows list, for various species, which genomic markers are present. The hypothesis is that in a hypothetical ancestor these markers were consecutive. Therefore one is interested in deciding whether the columns of the matrix can be reordered to make nonzero entries in all rows consecutive. This property, known as *consecutive-ones-property* (C1P) can be checked in polynomial time(Fulkerson & Gross 1965; Tucker 1972), and is (up-)monotone under taking submatrices. If the matrix fails to be C1P, one needs to find out which rows are the cause: The minimal conflicting sets of rows are determined by the dual monotone function (Chauve *et al.* 2009).

Finally, some applications give rise to boolean functions that are not monotone, but where maximal solutions are of interest: As introduced in (Saez-Rodriguez *et al.* 2007), a structured class of $\mathcal{NP}$-complete satisfiability problems is well-suited to model the logical behavior of cellular signaling processes. The resulting SAT problems are often infeasible, for various reasons, and one needs to compute all maximal feasible submodels under a certain structured relaxation (Haus *et al.* 2009). This reformulation does not monotonously relax the initial problem, but it is well-known that a monotone relaxation can be computed that preserves the maximal solutions (Kavvadias, Sideri, & Stavropoulos 2000). While, in general, this relaxation may exponentially increase formula size, we could show that in this case it actually always significantly decreases formula size. The original question could therefore be called 'almost-motonone', since its monotone core becomes apparent under the reformulation. In this paper we propose another 'almost-monotone' property, where the intersection of two monotone functions is used to analyze biological data.

Given a set of objects and a set of characteristics that each object might exhibit or not, the task of *frequent itemset mining* is to find subsets of characteristics that are exhibited by large subsets of objects. This way, the objects are gouped w.r.t. shared characteristics. An application of this principle can be found in the analysis of shopping baskets in supermarkets. The task is to find out which items are frequently bought together. A shopping basket is identified with the items it contains, i.e., $B = \{\text{milk}, \text{bread}, \text{juice}\}$. There is a finite number of items that might be contained in a shopping basket. This reasoning gives rise to a 0/1-matrix with the rows corresponding to the shopping baskets analyzed and the columns corresponding to the items that might be bought and thus contained in the individual baskets. Frequent itemset mining is a well-studied area in the field of datamining.(Goethals 2005; Uno, Kiyomi, & Arimura 2005; Boros *et al.* 2003) The focus mostly lies on the computation of all *maximal* frequent itemsets with a prescribed lower bound for the frequency where attention lies on high bounds, often expressed as 90% of the total number of rows.

In our application to the analysis of biological data however, the task is slightly different. We want to identify subsets of properties characterizing subsets of objects and at the same time separate them from all the other objects. The subsets of properties do not necessarily have to be maximal. In biological applications, it is often advantageous if these sets of charcaterizing properties are small: each property corresponds to a measurement that has to be carried out. In order to minimize the number of measurements, the characterizing sets of properties should be rather small. On the other hand, large property sets carry the information that many ofthe properties belong together. We want to find all possibilities to group the objects in this way.

Furthermore, in our applications there are usually only few objects and a large number of characteristics that might or might not be exhibited by the objects. This is due to the fact that in biological applications there is usually only a small number of subjects. E.g., in neuroscience the subject needs to be trained, the number of specimen is small etc. A huge variety of measurements is carried out with respect to these subjects: these give rise to the eventually characterizing properties of the subjets. As the number of possibly characterizing properties is large compared to the number of studied subjects, in most cases there are many different ways to group the subjects according to similarities of certain subsets of characteristics. As we want to find out all the ways to group the subjects in this way, an enumerative approach seems necessary. In this paper, we will point out certain aspects that will facilitate and speed up such an enumerative approach. Furthermore, we will present a way of computing groupings of the objects that are uniform in the sense that the groups stay relatively stable even when only subsets of their characterizing properties are considered.

2 Frequent and infrequent sets

In our biological application, the subjects that are studied give rise to the rows of a matrix. The possibly characterizing features give rise to the columns of this matrix. In this way, a 0/1-matrix arises, where a 1-entry implies that the subject corresponding to the row exhibits the feature corresponding to the column of the matrix-entry. Therefore, let $\mathbf{A} \in \{0,1\}^{m\times n}$ be a binary matrix. For $I \subseteq \{1,\dots,m\}$ let $\mathbf{A}^I$ be the matrix consisting of the rows index by elements contained in the index set I. Analogously for $J \subseteq \{1,\dots,n\}$ let $\mathbf{A}_J$ be the matrix consisting of the columns indexed by elements contained in J.

The overall aim is to find column subsets $J \subseteq \{1,\dots,n\}$ characterizing certain rows of matrix $\mathbf{A}$ in a uniform way. Uniformity means that if J characterizes a certain set of rows of matrix $\mathbf{A}$ then also every subset of J characterizes more or less the same set of rows. We say that a set of columns J "characterizes" a set of rows I, if $\mathbf{A}^I_J = \mathbf{1}$, i.e., if every entry in the matrix $\mathbf{A}^I_J$ is equal to 1. If the set I has cardinality t, then the set J is called

I	$\begin{matrix}1 & 1 & 1 & \dots & 1\\ \vdots & \vdots & \vdots & \vdots & \vdots\\ 1 & 1 & 1 & \dots & 1\end{matrix}$	$\left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\}t_1$ rows
II	$\begin{matrix}0 & 1 & 1 & \dots & 1\\ \vdots & \vdots & \vdots & \vdots & \vdots\\ 0 & 1 & 1 & \dots & 1\end{matrix}$	$\left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\}t_2-t_1$ rows
	$\begin{matrix}1 & 0 & 1 & \dots & 1\\ \vdots & \vdots & \vdots & \vdots & \vdots\\ 1 & 0 & 1 & \dots & 1\end{matrix}$	$\left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\}t_2-t_1$ rows
	$\begin{matrix}\vdots & \vdots & \vdots & \dots & \vdots\end{matrix}$	
	$\begin{matrix}1 & 1 & 1 & \dots & 0\\ \vdots & \vdots & \vdots & \vdots & \vdots\\ 1 & 1 & 1 & \dots & 0\end{matrix}$	$\left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\}t_2-t_1$ rows
III	$\begin{matrix}* & * & * & \dots & *\\ \vdots & \vdots & \vdots & \vdots & \vdots\\ * & * & * & \dots & *\end{matrix}$	

Table 1: 3-block structure of the matrix $\mathbf{A}_J$

t-frequent as it characterizes at least t rows of matrix $\mathbf{A}$. A formal definition is given in Definition 1.

Definition 1. Let $\mathbf{A} \in \{0,1\}^{m\times n}$ and $t \in \mathbb{Z}_+$. An index set $J \subseteq \{1,\dots,n\}$ is called *t-frequent* if there exists an index set $I \subseteq \{1,\dots,m\}$ with $|I| = t$ such that $\mathbf{A}^I_J = \mathbf{1}$, i.e., if all entries of $\mathbf{A}^I_J$ are equal to 1. The set J is called *t-infrequent* otherwise. An index set J is called *maximally t-frequent* if it is t-frequent and if every proper superset is t-infreqent. J is called *minimally t-infrequent* if it is t-infrequent and if every proper subset is t-frequent.

t-frequent sets of columns J select a subset of rows I but do not specify anything about $\{1,\dots,m\} \setminus I$. To alleviate this, we propose to use a stronger notion:

Definition 2 ((t_1,t_2)-FMI set)**.** For given parameters $t_1, t_2 \in \{1,\dots,m\}$ with $t_1 < t_2$ a set J will be called (t_1,t_2)-*FMI* if J is simultaneously t_1-frequent and minimally t_2-infrequent.

Of obvious interest is the case $t_2 = t_1 + 1$, but larger values of of t_2 are also useful: Allowing a larger distance k between t_1 and t_2 represents the question 'Characterize t_1 rows, and rule out $m - t_2$, while not caring about k rows'.

FMI-sets give rise to a 3-partition of the matrix $\mathbf{A}$:

Proposition 1 (Matrix 3-partition by FMI sets). Let J be such a (t_1,t_2)-FMI set. Then the rows of matrix $\mathbf{A}$ can be ordered in a way that matrix $\mathbf{A}_J$ takes the form proposed in Table 1.

The first block consists of t_1 rows with only 1 entries. This is due to the fact that J is t_1-frequent.

The second block consists of $|J| \cdot (t_2 - t_1)$ rows. For each index $j \in J$ there are $t_2 - t_1$ rows. These rows have a zero entry in the column corresponding to index j

and 1 entries otherwise. In particular, in the case of $t_2 = t_1 + 1$, the second block of matrix $\mathbf{A}_J$ consists of a complementary unit-matrix. The structure of block II is related to the minimal t_2-infrequency of column set J. As J is minimal t_2-infrequent, every proper subset is t_2-frequent. In particular, every subset of cardinality $|J| - 1$ is t_2-frequent.

In the following, we will require:

(a) If J is a (t_1, t_2)-FMI set, then there is no $\bar{t}_1 > t_1$ such that J is a $(\bar{t}_1, t_2)$-FMI set.

(b) If J is a (t_1, t_2)-FMI set, then there is no $\bar{t}_2 > t_2$ such that J is a $(t_1, \bar{t}_2)$-FMI set.

to hold. This way redundancy is avoided. FMI-sets respecting properties (*a*) and (*b*) will also be called *max-FMI-sets*.

Example 1. The following example illustrates the necessity of max-FMI-sets to avoid redundancy.

	a	b
I	1	1
	1	1
II	0	1
	1	0
III	0	1
	1	0
	0	1
	1	0
	1	1

Here, column set $\{a, b\}$ is a $(2, 3)$-FMI set. But clearly, $\{a, b\}$ is also a $(3, 5)$-FMI set.

Another reason why the concept of max-FMI-sets reduces redundancy is given in the following lemma: (t_1, t_2)-FMI sets J of cardinality 1 are (t_1, t)-FMI for all $m \geq t \geq t_2$.

Lemma 1. Let $J \subseteq \{1, \ldots, n\}$ with $|J| = 1$. Suppose J is t-infrequent. Then J is minimally t'-infrequent for all $t' \geq t$.

Proof. As $|J| = 1$, its only proper subset is the empty set: $\emptyset$. The empty set is t-frequent for all $t \in \{1, \ldots, m\}$. As J is t-infrequent, the property of its only proper subset being t-frequent yields that J is minimally t-infrequent. An analogous argument yields that J is minimally t'-infrequent for all $t' \geq t$, $t' \in \{1, \ldots, m\}$. □

FMI-ness is a hereditary property. Now we will show that every proper subset of a (t_1, t_2)-FMI set is again a (s_1, s_2)-FMI set for some $s_1, s_2 \in \mathbb{Z}_+$ with $s_2 > s_1 \geq t_2$.

Lemma 2. Let $t_1, t_2 \in \mathbb{Z}_+$ and $t_1 < t_2$. Furthermore let $\mathbf{A} \in \{0, 1\}^{m \times n}$ and let $J \subseteq \{1, \ldots n\}$ be a (t_1, t_2)-FMI set. Let $\bar{J} \subsetneq J$ be a proper subset of J. Then there exist parameters $s_1, s_2 \in \mathbb{Z}_+$ with $s_2 - s_1 \geq t_2 - t_1$ and $s_1 \geq t_1 + |J \setminus \bar{J}| \cdot (t_2 - t_1)$ such that $\bar{J}$ is a (s_1, s_2)-FMI set.

Proof. Let $\bar{J} \subseteq J$. When considering the representation of matrix $\mathbf{A}_J$ proposed in Table 1 it is clear that $(J \setminus \bar{J}) \cdot (t_2 - t_1)$ rows from block II may be moved to block I when retsricting our attention to $maA_{\bar{J}}$. Therefore, $\bar{J}$ is s_1-frequent for $s_1 \geq t_1 + |J \setminus \bar{J}| \cdot (t_2 - t_1)$. It remains to prove that $\bar{J}$ is minimally s_2-infrequent for some s_2.

Let us assume first that $|\bar{J}| \geq 2$. Then our claim follows directly from the structure of the matrix $\mathbf{A}_J$ described in Table 1. For $\mathbf{A}_{\bar{J}}$, rows from the third block could possibly be moved to the first block. The second block however still consists of $|\bar{J}|(t_2 - t_1)$ rows that are all-one except for one 0 entry in each row. For every entry $j \in \bar{J}$ there are exactly $t_2 - t_1$ rows with a 0-entry in this block. Therefore, $\bar{J}$ is minimally s_2-infrequent for some $s_2 \geq s_1 + (t_2 - t_1) \geq s_1$.

Now let us assume that $|\bar{J}| = 1$. Then there are two cases: either $\mathbf{A}_{\bar{J}}$ is the all-one vector. Then $s_1 = m$ and $\bar{J}$ is minimally $m + 1$-infrequent. (Note that this is a quite artificial situation though.) Or $\mathbf{A}_{\bar{J}}$ is not the all-one vector. Then $\bar{J}$ s_1-frequent with $s_1 = \text{supp}(\mathbf{A}_{\bar{J}})$. Furthermore it is minimally $s_1 + 1$-infrequent. □

Uniformity of (t_1, t_2)-FMI sets. In our applications, we want to find sets of features (corresponding to columns of the binary matrix $\mathbf{A}$) that characterize certain subsets of subjects (i.e. rows of the matrix $\mathbf{A}$). It is essential that this notion of characterization stays stable when taking subsets of thecharacterizing features. In particular, we prefer sets of features where leaving out a single feature does not drastically change the set of characterized subjects.

To model this, we will assign an integer value to every FMI-set $J \subseteq \{1, \ldots, n\}$ Small values will correspond to sets that are in a sense *uniformly* FMI.

The value $\text{val}(J)$ of a set J is defined recursively. Let J be a (t_1, t_2)-FMI set. By Lemma 2 we know that every proper subset of J is a (s_1, s_2)-FMI set for some $s_1, s_2 \in \mathbb{Z}_+$ with $s_2 > s_1 \geq t_2$. We define a function $\sigma^J \colon J \to \mathbb{Z}_+^2$ which for $x \in J$ is returning the parameters (s_1, s_2) such that $J \setminus \{x\}$ is max-(s_1, s_2)-FMI. (In this notation, $\sigma_1^J(x) = s_1$ and $\sigma_2^J(x) = s_2$.) For all elements $i \in \{1, \ldots, n\}$ we set $\text{val}(\{i\}) := 0$. Furthermore for any set $J \subseteq \{1, \ldots, n\}$ that is (t_1, t_2)-FMI with $|J| \geq 2$ we set:

$$\text{val}(J) := \max_{x \in J} \{\sigma_1^J(x) - t_1,\ \text{val}(J \setminus \{x\})\}.$$

Example 2. In this example we want to illustrate how the value $\text{val}(J)$ of a set J is computed. Here, the column set J consists of three columns of a 0/1-matrix $\mathbf{A}$.

	a	*b*	*c*
I	1	1	1
II	0	1	1
	1	0	1
	1	1	0
III	1	1	0
	1	1	0
	1	1	0

Let $J := \{a, b, c\}$. Clearly, J is $(1, 2)$-FMI. Furthermore $\text{val}(J) = \sigma_1^J(c) - 1 = 5 - 1 = 4$.

The value of a set J is thus determined via the analysis of all its subsets. This could require $2^{|J|}$ computations of such values. The following lemma shows that it is sufficient to consider all pairs of elements contained in J to determine its overall value.

Lemma 3. Let $\mathbf{A} \in \{0,1\}^{m \times n}$ be a binary matrix and let $J \subseteq \{1, \ldots, n\}$ be (t_1, t_2)-FMI. Then

$$\operatorname{val}(J) = \max_{\substack{i,j \in J, \\ i \neq j}} \{\operatorname{val}(\{i,j\})\}. \tag{1}$$

Proof. Assume not. Then there is $I \subseteq J$ with $|I| \geq 2$ and (s_1, s_2)-FMI and there is $x \in I$ such that

$$\operatorname{val}(J) = \operatorname{val}(I) = \sigma_1^I(x) - s_1$$

This means there are $\sigma_1^I(x)$ rows, denoted by Z^x, that are all-one when restricted to the index set $I \setminus \{x\}$. In particular, only $\sigma_1^I(x)$ of these rows, denoted by Z, are all-one when restricted to I. Therefore, the column corresponding to the index x is zero on the rows $Z^x \setminus Z$. Let $y \in I \setminus \{x\}$ be arbitrary. Then y is all-one on all rows contained in $Z^x \setminus Z$. Therefore

$$\operatorname{val}(J) \geq \operatorname{val}(\{x,y\}) \geq |Z^x \setminus Z| = \sigma_1^I - s_1 = \operatorname{val}(J).$$

Therefore, $\operatorname{val}(\{x,y\}) = \operatorname{val}(J)$. This shows that the pair $\{x,y\} \subseteq J$ realizes the value of column set J. □

To determine the value of J it suffices thus to determine the values of all pairs of elements contained in J. This can be done with $(|J|-1)^2$ many operations.

Are pairs of columns enough? Lemma 3 reduces the computation of $\operatorname{val}(J)$ to polynomially many subproblems by considering only pairs of columns. Note however, that this does not imply that only pairs of columns make up for intersting FMI-sets. If we want to select a given subset of rows I using a set of columns J, then $|J|$ may need to be greater than 2. This is illsutrated in the following example.

1	1	1	1	rows we want
1	1	1	1	to characterize by
1	1	1	1	subsets of the columns
1	1	0	0	
1	1	0	1	
1	1	1	0	
1	0	1	1	
0	1	1	1	
0	0	0	0	

pairs are not enough!

3 Computational Results

We believe that the proposed method of studying FMI sets will support thorough analysis of data sets that are too small to be used in methods that rely on statistical reliability estimates. Such data is of increasing interest:

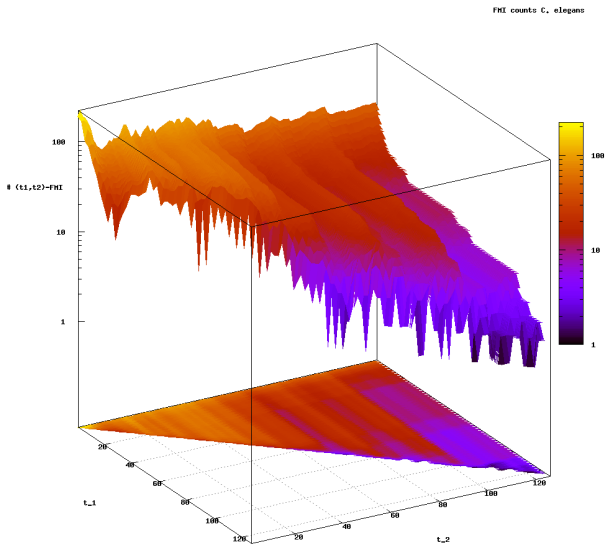

Figure 1: FMI sets in *C. elegans* data

modeling of rare events, expensive biological experiments, or long data acquisition processes may make high-volume data analysis unfavorable.

We will employ FMI sets to study a data set describing the association of gene defects with phenotypes of *Cenorhabdis elegans*, a nematode that is a well-established model organism in epigenetic research, being the first multicellular organism for which the full genome was sequenced, and RNAi experiments are easy to do. The data used is a table associating 35 (binary) patterns of gene defects on 9 genes and two types of treatments to 6 observed phenotypes. Both treatments and phenotypes are nonbinary (but discrete), and have been encoded in a unary fashion. The resulting matrix has 127 rows and 49 columns.

The first goal is to find all sets of genes that characterize a certain group of patterns, so that they are

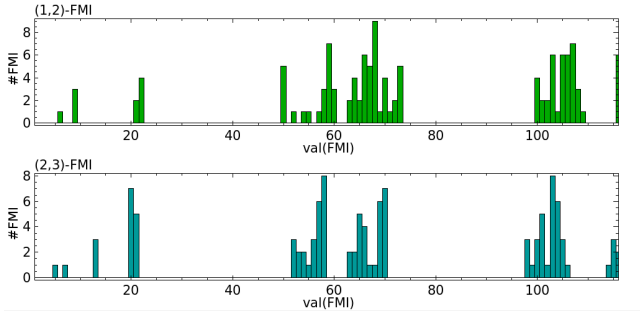

Figure 2: Histogram of values for $(1,2)$-FMI and $(2,3)$-FMI sets.

well distinguished from the complement of the group. For this purpose we computed all (t_1, t_2)-FMI sets, as shown in Figure 1. The maximum is reached at 221 sets for the group of $(1, 3)$-FMI sets, but as can be seen from the logarithmic scale the sizes are mostly below 60. Computations were performed using the publicly available implementation of the Joint Generation algorithm by the second author(Haus 2008). Since we need to compute both frequent and infrequent sets to compute FMI sets using a joint generation method is natural; one could obviously also use a specialized frequent set computation code and then compute the transversal of the result using e.g. Berge's method.

It is easy to compute values for the FMI sets using Lemma 3. The values vary widely, but appear clustered in three groups: small, medium and large values, as shown in the exemplary histograms in Figure 2.

Detailed analysis of the associations is still unfinished, and will be the focus of biological investigations, but initial results are very promising: The FMI sets with smallest values containing pathological phenotypes in fact reproduce the well-known association of unc-73 deficiency and impaired cell migration as a $(2, 3)$-FMI, of unc-73 deficiency and impaired axon guidance as another $(2, 3)$-FMI, and the association of mig-2 mutation with deficient cell migration.

Acknowledgement The first author was funded by a PAKT project of the German Leibniz Society. The second author was funded by the Magdeburg Center for Systems Biology (MaCS) and the state of Saxony-Anhalt. The third and fourth author were funded by the Freiburg Initiative in Systems Biology (FRISYS). MaCS and FRISYS are FORSYS centers financed by the German Ministry for Education and Research.

References

[Boros *et al.* 2003] Boros, E.; Gurvich, V.; Khachiyan, L.; and Makino, K. 2003. On maximal frequent and minimal infrequent sets in binary matrices. *Annals of Mathematics and Artificial Intelligence* 39(3):211–221.

[Chauve *et al.* 2009] Chauve, C.; Haus, U.-U.; Stephen, T.; and You, V. P. 2009. Minimal conflicting sets for the consecutive ones property in ancestral genome reconstruction. In *Comparative Genomics. Proceedings of International Workshop RECOMB-CG 2009, Budapest, Hungary, September 27-29, 2009*, volume 5817 of *Lecture Notes in Bioinformatics*, 48–58. Springer.

[Fulkerson & Gross 1965] Fulkerson, D. R., and Gross, O. A. 1965. Incidence matrices and interval graphs. *Pacific Journal of Mathematics* 15:835–855.

[Goethals 2005] Goethals, B. 2005. Frequent set mining. In Maimon, O., and Rokach, L., eds., *The Data Mining and Knowledge Discovery Handbook*. Springer. 377–397.

[Haus *et al.* 2009] Haus, U.-U.; Niermann, K.; Truemper, K.; and Weismantel, R. 2009. Logic integer programming models for signaling networks. *Journal of Computational Biology* 16(5):725–743.

[Haus, Klamt, & Stephen 2008] Haus, U.-U.; Klamt, S.; and Stephen, T. 2008. Computing knockout strategies in metabolic networks. *Journal of Computational Biology* 15:259–268.

[Haus 2008] Haus, U.-U. 2008. `cl-jointgen`, a Common Lisp Implementation of the Joint-Generation Method. available at `http://www.primaldual.de/cl-jointgen`.

[Kavvadias, Sideri, & Stavropoulos 2000] Kavvadias, D. J.; Sideri, M.; and Stavropoulos, E. C. 2000. Generating all maximal models of a boolean expression. *Information Processing Letters* 74(3-4):157–162.

[Klamt, Haus, & Theis 2009] Klamt, S.; Haus, U.-U.; and Theis, F. 2009. Hypergraphs and cellular networks. *PLoS Computational Biology* 5(5):e1000385.

[Saez-Rodriguez *et al.* 2007] Saez-Rodriguez, J.; Simeoni, L.; Lindquist, J.; Hemenway, R.; Bommhardt, U.; Arndt, B.; Haus, U.-U.; Weismantel, R.; Gilles, E. D.; Klamt, S.; and Schraven, B. 2007. An extensive logical model predicts key events in the T-cell receptor signaling network. *PLoS Computational Biology*. DOI `http://dx.doi.org/10.1371/journal.pcbi.0030163`.

[Tucker 1972] Tucker, A. C. 1972. A structure theorem for the consecutive 1's property. *Journal of Combinatorial Theory B* 12:153–162.

[Uno, Kiyomi, & Arimura 2005] Uno, T.; Kiyomi, M.; and Arimura, H. 2005. LCM ver. 2: Efficient mining algorithms for frequent/closed/maximal itemsets. In Bayardo, R.; Goethals, B.; and Zaki, M. J., eds., *Proceedings of the IEEE ICDM Workshop on Frequent Itemset Mining Implementations Brighton, UK, November 1, 2004*, volume 126 of *CEUR Workshop Proceedings*.