①

Algebra across the borders

August 8, 2011

Algorithms that produce or count all closed sets

Marcel Wild

A) Ganter's 1984 algorithm

B) The implication n-algorithm

c) Binary decision diagrams

A) Ganter's 1984 algorithm

Let $A \mapsto \overline{A}$ be any explicitely given closure operator on a finite set W. For instance, if $\mathcal{C} \subseteq 2^W$ is the closure system of all closed sets, and the system $\mathcal{M} \subseteq \mathcal{C}$ of its meet-irreducible members is known, then $\overline{A}$ can be computed as

$$\overline{A} = \bigcap \{ X \in \mathcal{M} : X \supseteq A \}.$$

Alternatively, if an implicational base Σ of $\mathcal{C}$ is known, then

$$\overline{A} = A' \cup A'' \cup A''' \cup \cdots ,$$

where generally

$$B' := B \cup \bigcup \{ V_i : U_i \to V_i \text{ in } \Sigma,\ U_i \subseteq B \}$$

Ganter's algorithm, sometimes called <u>Nextclosure</u>, generates all members of $\mathcal{C}$ in lexicographic order.

To fix ideas, take $W = \{1,2,3,4,5,6\}$ and <u>define</u> the closure operator $A \mapsto \overline{A}$ by the implicational base

$$\Sigma := \{\{1,2,3\} \to \{5,6\},\ \{4,5\} \to \{3,6\}\}.$$

Starting with $\phi = (0,0,0,0,0,0)$, the lexicographically first few sets happen to be Σ-closed:

1	2	3	4	5	6	
0	0	0	0	0	0	
0	0	0	0	0	1	
0	0	0	0	1	0	
0	0	0	0	1	1	
0	0	0	1	0	0	
0	0	0	1	0	1	$= \{4,6\}$ is closed
0	0	0	1	1	0	$= \{4,5\}$ is <u>not</u> closed

Let $X \subseteq W$ be closed. To find the <u>next</u> closed set, one processes the 0's of X (its char. vector) from the right:

1	2	3	4	5	6	
0	0	0	1	0	1	$= X$ is closed
0	0	0	1	1	0	$= X[5]$ (occupy 5th zero)
0	0	1	1	1	1	$= \overline{X[5]}$, fails
0	0	1	0	0	0	$= X[3]$ (occupy 3rd zero)
0	0	1	0	0	0	$\overline{X[3]}$ succeeds X
⋮	⋮	⋮	⋮	⋮	⋮	
1	1	0	1	0	1	$= Y$ is closed
1	1	0	1	1	0	$= Y[5]$
1	1	1	1	1	1	$= \overline{Y[5]}$ fails
1	1	1	0	0	0	$= Y[3]$
1	1	1	0	1	1	$= \overline{Y[3]}$ succeeds Y
⋮	⋮	⋮	⋮	⋮	⋮	

Proceeding like this one finds all 47 Σ-closed sets.

B) The implication n-algorithm

Consider again the implicational base

$$\Sigma = \{\{1,2,3\} \to \{5,6\}, \{4,5\} \to \{3,6\}\}$$

For the time being we focus on the first implication. The set Mod_1 of $(123 \to 56)$-closed sets can be represented as union of four $\{0,1,2\}$-valued rows:

	1	2	3	4	5	6
$r_1 =$	0	2	2	2	2	2
$r_2 =$	2	0	2	2	2	2
$r_3 =$	2	2	0	2	2	2
$r_4 =$	1	1	1	2	1	1

Notice that say $r_1 \cap r_2 \neq \emptyset$. That can be cured as follows:

	1	2	3	4	5	6
$r_1 =$	0	2	2	2	2	2
$r_2 =$	1	0	2	2	2	2
$r_3 =$	1	1	0	2	2	2
$r_4 =$	1	1	1	2	1	1

Now $Mod_1 = r_1 \dot\cup r_2 \dot\cup r_3 \dot\cup r_4$. From Mod_1 we sieve the family $Mod_2 \subseteq Mod_1$ of all sets X which are also $(45 \to 36)$-closed:

Mod_1

	1	2	3	4	5	6
r_1	0	2	2	0	2	2
	0	2	2	1	0	2
	0	2	2	1	1	2
r_2	1	0	2	0	2	2
	1	0	2	1	0	2
	1	0	2	1	1	2
r_3	1	1	0	0	2	2
	1	1	0	1	0	2
	1	1	0	1	1	2
r_4	1	1	1	2	1	1

Mod_2

1	2	3	4	5	6	
0	2	2	0	2	2	16
0	2	2	1	0	2	8
0	2	1	1	1	1	2
1	0	2	0	2	2	8
1	0	2	1	0	2	4
1	0	1	1	1	1	1
1	1	0	0	2	2	4
1	1	0	1	0	2	2
~~1~~	~~1~~	~~0~~	~~1~~	~~1~~	~~2~~	
1	1	1	2	1	1	2
						47

Consider now a more economic way based on the simple wildcard

$$\boxed{nn\cdots n := \text{"at least one zero here"}}$$

Recalling $\Sigma = \{123 \to 56,\ 45 \to 36\}$ we have:

	1	2	3	4	5	6
$r_1 =$	n	n	n	2	2	2
$r_2 =$	1	1	1	2	1	1

Mod_1

	1	2	3	4	5	6
r_1	n	n	n	$\bar{n}$	$\bar{n}$	2
	n	n	n	1	1	2
r_2	1	1	1	2	1	1

Mod_1

1	2	3	4	5	6	
n	n	n	$\bar{n}$	$\bar{n}$	2	42
n	n	1	1	1	1	3
1	1	1	2	1	1	2
						47

Mod_2

Here $42 = (2^3 - 1)\cdot(2^2 - 1)\cdot 2^1$.

A <u>Horn formula</u> $\varphi(x)$ is a conjunction of clauses each one of which with at most one positive literal, e.g.:

$$\varphi(x) = (\bar{x}_1 \vee \bar{x}_3 \vee x_5) \wedge (\bar{x}_4 \vee x_6) \wedge (\bar{x}_1 \vee \bar{x}_3 \vee x_7) \wedge x_2 \wedge x_9 \wedge (\bar{x}_1 \vee \bar{x}_6) \wedge (\bar{x}_2 \vee \bar{x}_5 \vee \bar{x}_8)$$

which is logically equivalent to

$$\underbrace{(x_1 \wedge x_3 \to x_5 \wedge x_7) \wedge (x_4 \to x_6) \wedge (T \to x_2 \wedge x_9)}_{=:\,\Sigma}$$

$$\wedge \underbrace{(\bar{x}_1 \vee \bar{x}_6) \wedge (\bar{x}_2 \vee \bar{x}_5 \vee \bar{x}_8)}_{=:\,\ominus}$$

We prefer set notation:

$$\Sigma = \{\{1,3\} \to \{5,7\}, \{4\} \to \{6\}, \phi \to \{2,9\}\}$$

$$\ominus = \{\{1,6\}^*, \{2,5,8\}^*\}$$

As usual, $x \in \{0,1\}^9$ is a model of φ if $\varphi(x)=1$. The models of φ match those subsets $X \subseteq \{1,2,\dots,9\}$ which are simultaneously $\leq$-closed (usual sense) and $\ominus$-noncovers, meaning that $X \not\supseteq \{1,6\}$ and $X \not\supseteq \{2,5,8\}$.

The next Theorem rests on these two key ingredients:

(a) By suitable precautions one can prevent that the implication n-algorithm encounters a situation where a $\{0,1,2,n\}$-valued row needs to be deleted.

(b) There is an efficient recursive way to count the k-element sets in a row of type

$0 \cdots 0$	$1 1 \cdots 1$	$2 \cdots 2$	$n_1 n_1 \cdots n_1$	$n_2 n_2 \cdots n_2$	$\cdots$	$n_t n_t \cdots n_t$

<u>**Thm**</u>: Consider a Horn formula, given as a family $\Sigma \cup \Theta$ of implications and subsets of W. Put $v := |\Sigma \cup \Theta|$, $w := |W|$.

(i) The number N of models $X \subseteq W$ can be found in time $O(Nv^2w^2)$ (in fact $O(Rv^2w^2)$).

(ii) Fix $k \le w$. The number N of models X with $|X| \le k$ can be found in $O(Nk^3v^2w^2)$.

(iii) Suppose the number of k-element models increases as k increases. Then the number N of models X with $|X| = k$ can be found in $O(Nk^4v^2w^2)$.

(iv) Suppose $\Sigma = \emptyset$ and $v := |\Theta|$ is $\le w$. For fixed $k \le w - v$ the number N of models X with $|X| = k$ costs $O(Nk^3v^2w^2)$.

(v) Fix $k \le w$, and no extra assumptions. The number N of models X with $|X| = k$ can be found in $O(N2^v w^6)$.
(The naive way costs $O(w^{k+2})$. If say $v = k = w$ then $2^v / w^{k+2} \longrightarrow 0$ as $w \longrightarrow \infty$.)