Verification of some Boolean partial polymorphisms

Mike Behrisch[*][†]

30th November 2021

Notation and basic definitions

Throughout we consider *partial Boolean operations*, that is, maps $f\colon D \to A$ defined on a subset $D \subseteq A^n$ for some $n \in \mathbb{N}$, where $A = \{0, 1\}$. The subset D where f is defined is also called the *domain* of f and denoted as $D = \operatorname{dom}(f)$; the integer n is the *arity* of f. A *Boolean relation* is any subset $\rho \subseteq A^m$ for some $m \in \mathbb{N}$, called the *arity* of ρ. In other words, an m-ary Boolean relation is any (possibly empty) set of m-tuples $(x_1, \dots, x_m)$ over $A = \{0, 1\}$.

Definition. *Let $m, n \in \mathbb{N}$. We say that an n-ary partial (Boolean) operation f* preserves *an m-ary (Boolean) relation ρ if the following condition, denoted by $f \rhd \rho$, holds: for every $(m \times n)$-matrix*

$$X = \begin{pmatrix} x_{11} & \dots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{m1} & \dots & x_{mn} \end{pmatrix} \in A^{m\times n}$$

with the property that all its rows belong to the domain of f,

$$(x_{i1} \dots x_{in}) \in \operatorname{dom}(f), \qquad (1 \le i \le m)$$

and all its columns belong to the relation ρ

$$\begin{pmatrix} x_{1j} \\ \vdots \\ x_{mj} \end{pmatrix} \in \rho, \qquad (1 \le j \le n)$$

the resulting column when f is applied to all rows of X must again belong to ρ:

$$f(X) := \begin{pmatrix} f(x_{11}, \dots, x_{1n}) \\ \vdots \\ f(x_{m1}, \dots, x_{mn}) \end{pmatrix} \in \rho.$$

[*]Institute of Discrete Mathematics and Geometry, TU Wien, Wien, Austria https://orcid.org/0000-0003-0050-8085

[†]Institute for Algebra, JKU Linz, Linz, Austria

Rephrasing the previous definition, a partial operation $f\colon D \to A$ fails to preserve $\rho \subseteq A^m$ if *there is* a matrix $X \in A^{m\times n}$ that acts as a counterexample, that is, all rows of X are in $\operatorname{dom}(f)$, all columns of X are in ρ, but $f(X) \notin \rho$. This means that non-preservation can be proven by a certificate of satisfiability of a first-order formula with $m \cdot n$ variables (one for each entry of the matrix $X \in A^{m\times n}$) expressing the relational constraints on the rows, the columns and the image of the matrix under f. The other way round, preservation can be justified by a proof of unsatisfiability of this formula. Both are tasks that can be handled efficiently by sat solvers, in particular in the Boolean case.

We observe also that $f\colon D \to A$ preserves a relation $\rho \subseteq A^m$ for trivial reasons if all matrices with columns from ρ have at least one row outside the domain D of f, or all matrices made up from rows of D have at least one column that does not belong to ρ, because then the corresponding first-order formula is clearly unsatisfiable.

If Q is a set of relations of possibly different arity, then we put

$$\operatorname{pPol} Q := \bigcup_{n\in\mathbb{N}} \bigcup_{D\subseteq A^n} \{\, f\colon D \to A \mid \forall \rho \in Q\colon f \rhd \rho \}\,,$$

and call this the *set of all partial polymorphisms of* Q.

Furthermore, as basic binary Boolean operations we need Boolean conjunction (and) $\wedge$ and addition modulo two (xor) $\oplus$.

Description of the dataset

The purpose of this dataset is to give a formal verification that the following partial ternary Boolean function

$$f\colon D \to \{0,1\}; \qquad f(x,y,z) := x \wedge y \wedge z$$

for all $(x,y,z) \in D \subseteq \{0,1\}^3$, that is,

$$f(x,y,z) = \begin{cases} 1 & \text{if } x = y = z = 1,\\ 0 & \text{for any other } (x,y,z) \in D, \end{cases}$$

where

$$D = \left\{ \begin{array}{c} (0,0,0),\\ (1,0,0),\\ (1,0,1),\\ (1,1,0),\\ (1,1,1) \end{array} \right\},$$

is a partial polymorphism of certain Boolean relations, but does not preserve certain others.

The involved relations are the following; we present them by a set-theoretic description as well as by listing all their elements by showing a Boolean matrix the columns of which are exactly the tuples in the relation. The meaning of the notation for these relations is relevant in a different context, but can be safely ignored here.

$$\Gamma_{\mathsf{L}_0}(\chi_2) = \{0\} \times \mathsf{ev}_3 = \left\{ (x_0,x_1,x_2,x_3) \in \{0,1\}^4 \;\middle|\; x_0 = 0 \;\&\; x_1 \oplus x_2 \oplus x_3 = 0 \right\}$$

$$
\begin{aligned}
&= \begin{Bmatrix} 0\,0\,0\,0 \\ 0\,1\,1\,0 \\ 1\,0\,1\,0 \\ 1\,1\,0\,0 \end{Bmatrix}, \\
\Gamma_{\mathsf{L}_2}(\chi_3) &= \left\{ (x_0,\dots,x_7) \in \{0,1\}^8 \;\middle|\; \begin{array}{r} (\exists i \in \{0,1,2\}\,\forall b_0,b_1,b_2 \in \{0,1\} \colon x_{4b_2+2b_1+b_0} = b_i) \\ \vee\, \forall b_0,b_1,b_2 \in \{0,1\} \colon x_{4b_2+2b_1+b_0} = b_2 \oplus b_1 \oplus b_0 \end{array} \right\}, \\
&= \begin{Bmatrix} 0\,0\,0\,0 \\ 0\,0\,1\,1 \\ 0\,1\,0\,1 \\ 0\,1\,1\,0 \\ 1\,0\,0\,1 \\ 1\,0\,1\,0 \\ 1\,1\,0\,0 \\ 1\,1\,1\,1 \end{Bmatrix}, \\
R_{\mathsf{L}} = \mathsf{ev}_4 &= \left\{ (x_1,x_2,x_3,x_4) \in \{0,1\}^4 \;\middle|\; x_1 \oplus x_2 \oplus x_3 \oplus x_4 = 0 \right\} \\
&= \begin{Bmatrix} 0\,1\,1\,1\,0\,0\,0\,1 \\ 0\,1\,0\,0\,1\,1\,0\,1 \\ 0\,0\,1\,0\,1\,0\,1\,1 \\ 0\,0\,0\,1\,0\,1\,1\,1 \end{Bmatrix}, \\
\{0\} \times R_{\mathsf{L}} = \{0\} \times \mathsf{ev}_4 &= \left\{ (x_0,x_1,x_2,x_3,x_4) \in \{0,1\}^5 \;\middle|\; x_0 = 0 \,\&\, x_1 \oplus x_2 \oplus x_3 \oplus x_4 = 0 \right\} \\
&= \begin{Bmatrix} 0\,0\,0\,0\,0\,0\,0\,0 \\ 0\,1\,1\,1\,0\,0\,0\,1 \\ 0\,1\,0\,0\,1\,1\,0\,1 \\ 0\,0\,1\,0\,1\,0\,1\,1 \\ 0\,0\,0\,1\,0\,1\,1\,1 \end{Bmatrix}, \\
\{0\} \times R_{\mathsf{L}} \times \{1\} = \{0\} \times \mathsf{ev}_4 \times \{1\} &= \left\{ (x_0,\dots,x_5) \in \{0,1\}^6 \;\middle|\; x_0 = 0 \,\&\, x_1 \oplus x_2 \oplus x_3 \oplus x_4 = 0 \,\&\, x_5 = 1 \right\} \\
&= \begin{Bmatrix} 0\,0\,0\,0\,0\,0\,0\,0 \\ 0\,1\,1\,1\,0\,0\,0\,1 \\ 0\,1\,0\,0\,1\,1\,0\,1 \\ 0\,0\,1\,0\,1\,0\,1\,1 \\ 0\,0\,0\,1\,0\,1\,1\,1 \\ 1\,1\,1\,1\,1\,1\,1\,1 \end{Bmatrix}.
\end{aligned}
$$

The aim of the dataset is to provide evidence for the following fact.

Claim. *The ternary Boolean conjunction f defined on D as above satisfies the following properties:*

$$f \in \mathrm{pPol}\,\{\Gamma_{\mathsf{L}_0}(\chi_2), \Gamma_{\mathsf{L}_2}(\chi_3)\}, \tag{1}$$
$$f \notin \mathrm{pPol}\,\{\{0\} \times R_{\mathsf{L}}\}, \tag{2}$$
$$f \notin \mathrm{pPol}\,\{\{0\} \times R_{\mathsf{L}} \times \{1\}\}. \tag{3}$$

Parts (2) and (3) can be quickly checked by hand, once a suitable matrix has been found. Such matrices are, for example,

$$X_2 = \begin{pmatrix} 0,0,0 \\ 1,0,0 \\ 1,0,1 \\ 1,1,0 \\ 1,1,1 \end{pmatrix} \overset{f}{\mapsto} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \notin \{0\} \times \mathsf{ev}_4 = \{0\} \times R_{\mathsf{L}},$$

consisting of the last three columns of $\{0\} \times R_{\mathsf{L}}$ and showing $f \not\triangleright \{0\} \times R_{\mathsf{L}}$, i.e. (2), and

$$X_3 = \begin{pmatrix} 0,0,0 \\ 1,0,0 \\ 1,0,1 \\ 1,1,0 \\ 1,1,1 \\ 1,1,1 \end{pmatrix} \overset{f}{\mapsto} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \notin \{0\} \times \mathsf{ev}_4 \times \{1\} = \{0\} \times R_{\mathsf{L}} \times \{1\},$$

consisting of the last three columns of $\{0\} \times R_{\mathsf{L}} \times \{1\}$ and showing that $f \not\triangleright \{0\} \times R_{\mathsf{L}} \times \{1\}$, i.e. (3).

The purpose of the present dataset is not to give these matrices, but to provide a script in the `SMT-LIB2.0` language that can be run by a sat solver such as Z3 [1, 2], which can perform this task automatically and, moreover, can prove unsatisfiability of the two corresponding problems related to (1). Such a script is given in the file `f-pPol-GammaL0chi2-GammaL2chi3.z3`. It makes use of the formalisation of non-preservation $(f \colon D \to \{0,1\}) \not\triangleright \rho$ as a satisfiability problem involving D, ρ and f. In this connection we represent all involved Boolean relations, such as D and various ρ, by their characteristic functions. Hence, the file `f-pPol-GammaL0chi2-GammaL2chi3.z3` only deals with Boolean operations; these are defined at the beginning of the script. The following table gives an overview of which function in `f-pPol-GammaL0chi2-GammaL2chi3.z3` represents which relation:

mathematical object	identifier used in `f-pPol-GammaL0chi2-GammaL2chi3.z3`
f	`f`
D	`domf`
$\{0\} \times \mathsf{ev}_3 = \Gamma_{\mathsf{L}_0}(\chi_2)$	`nev3`
$\{0\} \times \mathsf{ev}_4 = \{0\} \times R_{\mathsf{L}}$	`nev4`
$\{0\} \times \mathsf{ev}_4 \times \{1\} = \{0\} \times R_{\mathsf{L}} \times \{1\}$	`nev41`
$\Gamma_{\mathsf{L}_2}(\chi_3)$	`gL2chi3`

After defining these (characteristic) Boolean functions, for each of the four preservation problems in the claim, we first declare (existentially quantified) variables for the entries of the corresponding $(m \times 3)$-matrix using `declare-const`, and then we `assert` the constraints that have to hold for the rows, columns and the f-image of the matrix. Finally, we ask the solver to check whether this assertion is satisfiable (`check-sat`).

The output of running the Z3 solver on `f-pPol-GammaL0chi2-GammaL2chi3.z3` can be found in the file `z3-output.txt`. The files `f-preserves-GammaL0chi2_proof.txt` and `f-preserves-GammaL2chi3_proof.txt` contain formal proofs generated by Z3 corresponding to the preservation statements in (1).

References

[1] Leonardo de Moura and Nikolaj Bjørner. Z3: an efficient SMT solver. In Cartic R. Ramakrishnan and Jakob Rehof, editors, *International Conference on Tools and Algorithms for the Construction and Analysis of Systems (TACAS 2008)*, volume 4963 of *Lecture Notes in Comput. Sci.*, pages 337–340, Berlin, Heidelberg, March 2008. Springer. doi: `https://doi.org/10.1007/978-3-540-78800-3_24`.

[2] Microsoft Research. Z3 Theorem Prover, 2020. Available on-line from `https://github.com/z3prover/z3`.