A Classical Propositional Logic for Reasoning about Reversible Logic Circuits* **

Holger Bock Axelsen, Robert Glück, and Robin Kaarsgaard

DIKU, Department of Computer Science, University of Copenhagen,
Copenhagen, Denmark
{funkstar,glueck,robin}@di.ku.dk

Abstract. We propose a syntactic representation of reversible logic circuits in their entirety, based on Feynman's control interpretation of Toffoli's reversible gate set. A pair of interacting proof calculi for reasoning about these circuits is presented, based on classical propositional logic and monoidal structure, and a natural order-theoretic structure is developed, demonstrated equivalent to Boolean algebras, and extended categorically to form a sound and complete semantics for this system. We show that all strong equivalences of reversible logic circuits are provable in the system, derive an equivalent equational theory, and describe its main applications in the verication of both reversible circuits and template-based reversible circuit rewriting systems.

1 Introduction

Reversible computing – the study of computing models deterministic in both the forward and backward directions – is primarily motivated by a potential to reduce the power consumption of computing processes, but has also seen applications in topics such as static average-case program analysis [17], unified descriptions of parsers and pretty-printers [16], and quantum computing [6]. The potential energy reduction was first theorized by Rolf Landauer in the early 1960s [12], and has more recently seen experimental verification [2, 14]. Reaping these potential benefits in energy consumption, however, requires the use of a specialized gate set guaranteeing reversibility, when applied at the level of logic circuits.

Boolean logic circuits correspond immediately to propositions in classical propositional logic (CPL): This is done by identifying input lines with propositional atoms, and logic gates with propositional connectives, reducing the problem of reasoning about circuits to that of reasoning about arbitrary propositions in a classical setting. However, although Toffoli's gate set for reversible circuit logic is equivalent to the Boolean one in terms of what can be computed [22], it falls short of this immediate and pleasant correspondence. This article seeks

^{*} The authors acknowledge support from the Danish Council for Independent Research | Natural Sciences under the *Foundations of Reversible Computing* project, and partial support from COST Action IC1405 *Reversible Computation*.

^{**} Colors in electronic version.

to establish such a correspondence by proposing a standardized way of syntactically representing and reasoning about reversible logic circuits. This is done by considering a reformulation, and slight extension, of the toolset of classical propositional logic. The main contributions of this article are the following:

- A syntactic representation of entire reversible logic circuits as propositions, and a pair of proof calculi for reasoning about the semantics of thusly represented reversible logic circuits, sound and complete with respect to
- a categorical/algebraic semantics based on the free strict monoidal category over a *Toffoli lattice*, an order structure proven *equivalent to Boolean rings*,
- a proof that all strong equivalences of reversible logic circuits are provable,
 and
- an illustration of how the presented logic can be used to show strong equivalences of reversible circuits, and in particular to verify template-based reversible logic circuit rewriting systems.

The complexity of reversible circuits has been increasing while at the same time entirely new functional designs have been found (e.g. linear transforms [5], reversible microprocessors [21]). Established tools employing conventional Boolean logic are not geared towards the synthesis, transformation and verification of reversible circuits. Thus, it is important to find better ways of handling this new type of circuits, and some work has been approaching these problems from different angles (e.g. [4, 23]). Our goal is to formally model the semantics of reversible circuits, and in particular to capture strong equivalence of such circuits as provable equivalence of propositions.

Overview: Sec. 2 introduces the syntax and intuitive interpretation of the connectives, and shows how reversible logic circuits can be represented as propositions by way of a simple annotation algorithm. Sec. 3 describes the proof calculi used to reason about circuits thus represented, and relates them to existing systems. Sec. 4 develops the concept of a Toffoli lattice as a semantics for the central proof calculus and extends it, via a categorical view on such a structure, into the final model category \mathfrak{T}_{\otimes} . Sec. 5 sketches the fundamental metatheorems of soundness, completeness and circuit completeness, Sec. 6 outlines the applications of the developed theory in reversible circuit rewriting, and Sec. 7 presents ideas for future work, and concludes on the results presented.

2 Circuits as propositions

The correspondence between Boolean circuits and propositions, in all of its convenience to areas such as circuit design and computational complexity, did not happen by mistake: It is a well-known result that any Boolean function can be computed by a circuit composed of only NAND gates and constants, yet the Boolean gate set is still, in all of its redundancies, considered the *lingua franca* of logic circuit design, precisely due to this correspondence.

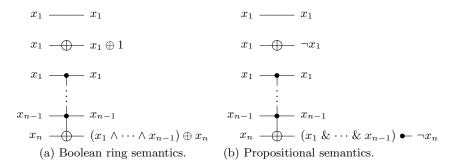


Fig. 1: Toffoli's reversible gate set – consisting of, from top to bottom, the IDENTITY gate, the NOT gate, and the generalized Toffoli gate – annotated with their Boolean ring semantics, as well as our propositional semantics.

Reversible circuits are usually depicted as gate networks where computation flows from left to right. Here, we consider circuits composed of the gates in Toffoli's reversible gate set, shown in Fig. 1a. (This widely used gate set is known as the *Multiple-Control Toffoli* (MCT) library.) We provide a brief exposition, which the reader familiar with reversible circuit logic can safely skip.

x_1	x_2	x_3		x_1	x_2	x_3
0	0	0		0	0	0
0	0	1		0	0	1
0	1	0		0	1	0
0	1	1	\mapsto	0	1	1
1	0	0		1	0	0
1	0	1		1	0	1
1	1	0		1	1	1
1	1	1		1	1	0

The only gate that warrants particular explanation is the generalized TOFFOLI gate, since the remaining gates

behave exactly as they do in Boolean circuit logic: This gate takes n>1 input lines, of which n-1 are control lines (marked with black dots), and the remaining one is the target line (marked with \oplus). If all control lines carry a value of 1, the value on the target line is negated – if not, the input of the target line simply passes through unchanged. As such, the control lines control whether the NOT operation should be carried out on the target line; in either case, the inputs to all control lines are carried through to the output unchanged (see also the truth table to the right for the generalized Toffoli gate where n=3; x_1 and x_2 are control lines, x_3 is the target line). Circuits may be composed horizontally (i.e., by ordinary function composition) and vertically (i.e., by computation in parallel) so long as they remain finite in size and contain neither loops, fan-in, nor fan-out. Note also that even though the target line is placed at the bottom in Fig. 1a for purposes of illustration, it may be placed anywhere relative to the control lines.

Contrary to Toffoli's Boolean ring semantics for the gate set [22], our presentation embraces Feynman's control interpretation [6] not just in the intuitive explanation given above, but also directly in the formalism. Following Kaarsgaard [11], this is done by replacing exclusive disjunction (here, $\cdot \oplus \cdot$) with the connective $\cdot \bullet - \cdot$, read as *control*, and introducing the usual negation connective $\neg \cdot$ on the target. This results in the propositional semantics shown in Fig. 1b. In this case, the semantics of the target line for the generalized TOFFOLI gate pleasingly reads as " x_1 and \cdots and x_{n-1} control not x_n ". While Soeken &

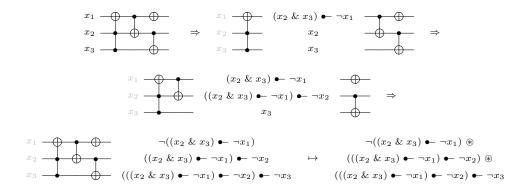


Fig. 2: An example of the annotation algorithm.

Thomsen [20] have shown (with their box rules) that control is a general concept corresponding (roughly) to conditional execution of a subcircuit, it turns out that, at the level of individual circuit lines, control carries the same meaning as material bi-implication in CPL. We postpone the proof of equivalence of these two approaches to Sec. 5.

Although the target line of the generalized TOFFOLI gate is, in many ways, the heart of this gate's semantics, it only paints part of the picture. Since reversible circuits are, by definition, required to have the same number of output lines as input lines, parallelism plays a much larger role in reversible circuits than in Boolean ones: To capture the semantics of reversible logic circuits in their entirety, we need a way to capture this parallelism. We do this by introducing yet another connective, $\cdot \circledast \cdot$, read as while, with the meaning of $A \circledast B$ as the multiplicative ordered conjunction of propositions A and B, i.e. as string concatenation in a free monoid. Order is important: as stated earlier, we wish the provable equivalence relation to capture strong equivalence of reversible circuits (reversible circuits are strongly equivalent if they compute the same function up to function extensionality [8]), rather than equivalence up to arbitrary permutation of output lines.

Using these two new connectives, along with the usual connectives for conjunction (here, $\cdot\&\cdot$) and negation, we can produce a straightforward annotation algorithm for extracting the semantics of reversible logic circuits as a proposition in this syntax. As also done for Boolean circuits, we identify each input line with a (fresh) propositional atom, and then propagate the semantics (as given in Fig. 1b) through until the entire circuit has been annotated, at which point we terminate and return the multiplicative ordered conjunction of these propositions, from top to bottom. An example of the annotation algorithm can be found in Fig. 2.

As also noted by Kaarsgaard [11], the syntax of propositions for forming reversible logic circuits using Toffoli's gate set in Fig. 1b is more restrictive than,

e.g., CPL; that is, (ordinary) conjunctions only appear as subpropositions of controls. Further, linear ordered conjunctions only appear as a way of "glueing" the propositions of individual circuit lines together (see, e.g., the final step in Fig. 2).

This structure suggests a syntactic hierarchy, which we will illustrate by means of color: Blue propositions will be those that correspond to the semantics of a single circuit line (perhaps of many in a circuit), red propositions correspond to the semantics of entire circuits (or subcircuits), and yellow (or recolorable) propositions will be those that can be either of these two. Formally, we define such propositions to be those produced by the grammars

$$\begin{array}{lll} A_B, B_B, C_B := A_Y \mid \neg A_B \mid A_B \& B_B & \text{(Blue propositions)} \\ A_Y, B_Y, C_Y := a \mid \mathbf{0} \mid \mathbf{1} \mid \neg A_Y \mid A_B \bullet - B_Y \mid A_Y \bullet - B_B & \text{(Yellow propositions)} \\ A_R, B_R, C_R := A_Y \mid A_R \circledast B_R \mid \mathbf{e} & \text{(Red propositions)} \end{array}$$

where a denotes any propositional atom; we will assume that there is a denumerable set P of these. For readability, we adopt the convention that $\neg \cdot$ binds tighter than $\cdot \& \cdot$, which binds tighter than $\cdot • - \cdot$, which finally binds tighter than $\cdot \circledast \cdot$. Further, we will omit subscripts when the syntactic class is clear from the context.

Starting with blue and recolorable propositions, $\mathbf{0}$ and $\mathbf{1}$ represent the false respectively true proposition (corresponding to ancillae, lines of constant value, in circuit terms), $\neg A_Y$ the usual negation of a proposition, $A_B \bullet - B_Y$ and $A_Y \bullet - B_B$ as "A control B", and finally $A_B \& B_B$ as the usual (additive) conjunction. Red propositions are interpreted as circuit structures, with $A_R \otimes B_R$ representing the ordered (parallel) structure made up of A_R and B_R , and \mathbf{e} representing the empty structure (i.e., the empty circuit). Further, we will denote the set of all such well-formed blue respectively red propositions by Φ_B respectively Φ_B .

In the same manner, well-formed blue and red contexts (a notion of a recolorable context is unnecessary) are those produced by the grammars

$$\begin{array}{ll} \varGamma_B, \Delta_B, \varPi_B := \cdot \mid \varGamma_B, A_B & \text{(Blue contexts)} \\ \varGamma_R, \Delta_R, \varPi_R := \cdot \mid \varGamma_R, A_R & \text{(Red contexts)} \end{array}$$

The distinction between the empty blue context and the empty red one is important, since the two types of contexts will be interpreted in two different ways; blue contexts are interpreted as an additive (blue) conjunction with $\mathbf{1}$ as unit, while red contexts are interpreted as an ordered multiplicative (red) conjunction with \mathbf{e} as unit. As we did for propositions, we will denote the set of all well-formed blue respectively red contexts by Φ_B^* respectively Φ_R^* .

3 Proof calculi

As the syntax presented in the previous section perhaps already alludes to, we will use not one but two proof calculi to reason about propositions thus formed. Figs. 3

Fig. 3: The blue fragment of LRS $_{\circledast}$.

and 4 show the two proof calculi – the blue and the red fragment, respectively – that make up the logic which we shall call LRS $_{\circledast}$.

There are two judgment forms, $\Gamma_B \vdash_B \varphi_B$ and $\Gamma_R \vdash_R \varphi_R$, which differ not only by syntax, but also by the interpretation of the context: Blue contexts are understood to be an additive (ordinary) conjunction of its constituent propositions (as usual) with **1** as unit, while red contexts are understood as a multiplicative ordered conjunction of its constituent propositions with **e** as unit. This difference of interpretation is reflected directly in the core rules of the calculi; while the identity and cut rules for the red fragment use careful bookkeeping to ensure that order and linearity are not broken, the corresponding rules in the blue fragment display implicit use of the structural rules available in the blue fragment. More explicitly, the blue fragment contains the usual structural rules of CPL – weakening, contraction, and exchange – while the red fragment has *none* of these.

The blue fragment, largely similar to the sequent calculus of LRS [11], presents itself as a reformulation of CPL in which control (corresponding to material bi-implication) is taken as a fundamental connective, while implication and disjunction are omitted. In particular, the omission of disjunction as a connective presents a challenge for classical reasoning, as we can no longer express the law of the excluded middle axiomatically in a way which facilitates its easy use in derivations. To resolve this, we present the rule instead as an explicit case analysis, corresponding to a proof tree of the form

Fig. 4: The red fragment of LRS $_{\circledast}$.

in CPL, which not only presents the common use case of the law of the excluded middle, but is also strong enough to derive the double negation elimination rule in the straightforward way. (Proof theoretically inclined buyers beware: Though this rule is sufficiently powerful, it threatens the subformula property [3] even in the face of cut-elimination.) Note that as we are not aiming for minimality, both rules are included in the blue fragment.

The red fragment offers little in terms of rules, since the only structure we are interested in is the parallel structure of circuit lines, captured by the rules for ordered multiplicative conjunction – essentially corresponding to concatenation of strings (though our formulation follows the conjunctive fragment of Polakow's presentation [15] of the Lambek calculus), with **e** corresponding to the empty string.

In our setting, by far the most interesting rule of the red fragment is the recoloring rule, which states that any logical deduction from a single recolorable proposition can be inserted into a structure of unit length, as long as the succedent is likewise recolorable. Recall that the recolorable propositions are precisely those that are well-formed as both blue and red propositions, so this (purely syntactic) side condition is entirely reasonable. Fig. 5 gives a larger example of an LRS_® derivation, showing $\neg x_1 \circledast x_1 \circledast \neg x_2 \vdash_R \neg x_1 \circledast \neg \neg x_1 • \neg x_2$.

Finally, it is worth noting that the syntax of red propositions is not strong enough to ensure that only reversible circuits can be represented. For example, the red proposition $x_1 \circledast x_1 \& x_2 - \neg x_3$ is perfectly well-formed, but does not correspond directly to a reversible circuit. On the other hand, every reversible circuit can adequately, and with minimal work, be represented as a red proposition, as we saw in Sec. 2. This turned out to result in an interesting tradeoff in the proof calculi: Naturally, it would be desirable if we could guarantee that every red proposition corresponded precisely to a reversible circuit – however, by not guaranteeing this property, we may consider the semantics of a single line or group of lines in isolation, without having to take the overall structure of the circuit into account at every step of a derivation, making for simpler overall logic.

4 Semantics

Given the obvious similarities between CPL and the blue fragment of LRS_{\circledast} , it would seem highly natural to adopt truth-functional semantics here as well. While this approach certainly works when considering the blue fragment in isolation, extending this approach to the red fragment runs into the problem of defining a single truth value – and for good reason, since truth should be interpreted relative to a circuit structure, taking order and resource use (*i.e.*, viewing circuit lines as ordered resources) into consideration.

For this reason, we will instead take the algebraic approach to semantics by considering what we call a $Toffoli\ lattice$, an order structure with obvious similarities to the blue fragment of LRS_{\circledast} . This approach has the immediate benefit that order structures can very easily be interpreted as categories, giving us a whole suite of tools to extend the semantics to the red fragment. We define Toffoli lattices, and their corresponding homomorphisms, as follows:

Definition 1. A Toffoli lattice $\mathfrak{A} = (A, \leq, \top, \bot, \wedge, \Leftarrow, \overline{\cdot})$ consists of a partially ordered set (A, \leq) furnished with the following operations and conditions:

- (i) There is a greatest element \top such that $x \leq \top$ for any element x.
- (ii) There is a least element \bot such that $\bot \le x$ for any element x.
- (iii) Given elements a, b there is an element $a \wedge b$ such that $x \leq a \wedge b$ iff $x \leq a$ and $x \leq b$.
- (iv) Given elements a, b there is an element a = b, the relative equivalence of a and b, such that $x \le a = b$ iff $x \land a \le b$ and $x \land b \le a$.
- (v) Given an element a, there is an element \overline{a} satisfying $x \leq \overline{a}$ iff $x \wedge a \leq \bot$, $a \wedge \overline{a} \leq \bot$, and if $x \wedge a \leq b$ and $x \wedge \overline{a} \leq b$ then $x \leq b$.

As is often done, we will use $|\mathfrak{A}|$ to denote the carrier set A.

Definition 2. Let \mathfrak{A} and \mathfrak{B} be Toffoli lattices. A Toffoli lattice homomorphism is a function $h: |\mathfrak{A}| \to |\mathfrak{B}|$ that preserves all lattice operations and constants, i.e., $h(\top) = \top$, $h(\bot) = \bot$, $h(a \land b) = h(a) \land h(b)$, $h(a \leftrightharpoons b) = h(a) \leftrightharpoons h(b)$, and $h(\overline{a}) = \overline{h(a)}$ for all $a, b \in |\mathfrak{A}|$.

From this definition, the truth-functional semantics appear by considering the set $\{0,1\}$:

Example 1. The set $\{0,1\}$ equipped with the usual partial order is a Toffoli lattice: Assigning the usual truth table semantics to \top , \bot , \wedge , and complement, and defining

$$0 \leftrightharpoons 0 = 1$$
 $0 \leftrightharpoons 1 = 0$ $1 \leftrightharpoons 0 = 0$ $1 \leftrightharpoons 1 = 1$

it is straightforward to verify that this yields a Toffoli lattice.

Though no explicit join operation is given, joins may be defined using meets and complements -i.e., analogously to Boolean algebras, one can show that $\overline{\overline{x} \wedge \overline{y}}$ is the least upper bound of x and y.

Lemma 1. Let $h: \mathfrak{A} \to \mathfrak{B}$ be a Toffoli lattice homomorphism. Then h is specifically monotonic, i.e. if $a \leq b$ then $h(a) \leq h(b)$.

Like so many other structured sets, these definitions lead us, without much trouble, to show that Toffoli lattices with homomorphisms between them form a concrete category; a useful feature which we will use shortly to characterize the free Toffoli lattice.

Theorem 1. The class of all Toffoli lattices with Toffoli lattice homomorphisms between them forms a category, TL.

Careful inspection of the definition of a Toffoli lattice reveals a correspondence with the blue fragment of LRS_{\circledast} – of course, this is entirely by design, though this correspondence is missing one part, namely the propositional atoms (recall the assumption that these form a denumerable set P). To account for these, we observe that Toffoli lattices may be freely constructed, and apply this free construction to the set of propositional atoms P to form an order theoretic model of the blue fragment.

Theorem 2. Toffoli lattices may be freely constructed, i.e., the forgetful functor $U: \mathbf{TL} \to \mathbf{Sets}$ has a left adjoint $F: \mathbf{Sets} \to \mathbf{TL}$.

Using this theorem, we take $\mathfrak{T} = FP$ (where P is the set of propositional atoms) to be our model of the blue fragment. This allows us to define blue denotation and entailment:

Definition 3. The denotation of a blue proposition $\varphi_B \in \Phi_B$, denoted $[\![\varphi_B]\!]$, is given by the function $[\![\cdot]\!]: \Phi \to |\mathfrak{T}|$ defined as follows:

where a denotes any propositional atom in P. Further, the denotation of a blue context $\Gamma_B \in \Phi_B^*$ is given by the overloaded function $\llbracket \cdot \rrbracket : \Phi_B^* \to |\mathfrak{T}|$ defined by

$$\llbracket \cdot \rrbracket = \top \qquad \qquad \llbracket \Gamma'_B, A_B \rrbracket = \llbracket \Gamma'_B \rrbracket \wedge \llbracket A_B \rrbracket$$

Definition 4 (Blue entailment). Let Γ be a well-formed blue context, and φ be a well-formed blue formula. Then we define the blue entailment relation by $\Gamma \vDash_B \varphi$ iff $\llbracket \Gamma \rrbracket \leq \llbracket \varphi \rrbracket$ in \mathfrak{T} .

In the same manner as for any other partially ordered set, we can regard a single Toffoli lattice $\mathfrak A$ as a (skeletal preorder) category by considering each element of $|\mathfrak A|$ as an object of the category, and placing a morphism between objects X and Y iff $X \leq Y$ in $\mathfrak A$. This allows us to extend our model lattice $\mathfrak T$ by categorical means to form a model of the red fragment. A key insight in this regard is the role of monoidal categories in modelling linear logic [18]; in particular, a strict monoidal category is sufficient to model the red fragment. This leads to the following construction:

Definition 5. Let \mathfrak{T}_{\otimes} denote the free strict monoidal category over \mathfrak{T} . That is, \mathfrak{T}_{\otimes} has as objects all strings $X_1X_2...X_n$ where all X_i are objects of \mathfrak{T} , and as morphisms all strings of morphisms $f_1f_2...f_n: X_1X_2...X_n \to Y_1Y_2...Y_n$ for morphisms $f_i: X_i \to Y_i$ of \mathfrak{T} . It has a monoidal tensor $-\otimes -: \mathfrak{T}_{\otimes} \times \mathfrak{T}_{\otimes} \to \mathfrak{T}_{\otimes}$ defined by concatenation; thus it is strictly associative and has a strict unit i, denoting the empty string.

See, e.g., Joyal & Street [9, 10] for the construction of the free strict monoidal category (or, in their nomenclature, free strict tensor category) over a given category \mathbf{C} ; it simply amounts to be the coproduct of all functor categories of the form \mathbf{C}^n , where n is the discrete category of n objects. This allows us to characterize \mathfrak{T}_{\otimes} by means of a (Grothendieck) fibration (see, e.g., Jacobs [7]) into the discrete category $\Delta_{\mathbb{N}}$ which has, as objects, all natural numbers¹:

Theorem 3. The functor $\Psi: \mathfrak{T}_{\otimes} \to \Delta_{\mathbb{N}}$ defined by mapping objects to their lengths as strings, and morphisms to the corresponding identities is a Grothendieck fibration and a monoidal functor. Specifically, each inverse image $\Psi^{-1}(k)$ for k in $(\Delta_{\mathbb{N}})_0$ is a full subcategory of \mathfrak{T}_{\otimes} .

Proof (*Proof sketch*). Since $\Delta_{\mathbb{N}}$ is discrete, for any object $X_1X_2...X_n$ of \mathfrak{T}_{\otimes} , the only possible morphism in $\Delta_{\mathbb{N}}$ of the form $u:K\to \Psi(X_1X_2...X_n)$ is the identity $1_{\Psi(X_1X_2...X_n)}$, which the identity morphism $1_{X_1X_2...X_n}$ is trivially cartesian over.

To see that Ψ is a strict monoidal functor, we note the obvious tensor product in $\Delta_{\mathbb{N}}$ given by addition, *i.e.*, by mapping objects $A \otimes B$ to their sum (as natural numbers) A + B, and likewise morphisms $1_A \otimes 1_B$ to 1_{A+B} . From this, it follows directly that $\Psi(A \otimes B) = \Psi(A) \otimes \Psi(B)$.

This approach is closely related to the theory of PROs, PROPs, and operads (see, e.g., Leinster [13]) – indeed, \mathfrak{T}_{\otimes} is a PRO – but we will avoid relying on this theory for the sake of a more coherent presentation.

In order to define denotation and entailment in the red propositions, we need one last lemma, stating the obvious isomorphism between $\Psi^{-1}(1)$ (the subcategory of strings of objects of \mathfrak{T} of length precisely 1) and \mathfrak{T} :

Lemma 2. There exist functors $I: \mathfrak{T} \to \Psi^{-1}(1)$ and $J: \Psi^{-1}(1) \to \mathfrak{T}$ witnessing $\Psi^{-1}(1) \cong \mathfrak{T}$.

Definition 6. The denotation of a red proposition $\varphi_R \in \Phi_R$, denoted $[\![\varphi_R]\!]$, is given by the function $[\![\cdot]\!]: \Phi_R \to (\mathfrak{T}_{\otimes})_0$ defined as follows:

$$\llbracket \mathbf{e} \rrbracket = i \quad \llbracket A_R \circledast B_R \rrbracket = \llbracket A_R \rrbracket \otimes \llbracket B_R \rrbracket \quad \llbracket A_R \rrbracket = I(\llbracket A_B \rrbracket) \quad \text{if } A_R \text{ is a recolorable.}$$

As we did for blue propositions, we overload the denotation function to apply to (in this case, red) contexts as well, by defining the function $\llbracket \cdot \rrbracket : \underline{\Phi}_R^* \to (\mathfrak{T}_{\otimes})_0$ as

$$\llbracket \cdot \rrbracket = i$$
 $\llbracket \varGamma_R, A_R \rrbracket = \llbracket \varGamma_R \rrbracket \otimes \llbracket A_R \rrbracket$

¹ We use the notation $\Delta_{\mathbb{N}}$ for the discrete category specifically to avoid confusion with the ordinal category ω , which some authors denote \mathbb{N} .

Definition 7 (Red entailment). Let Γ be a well-formed red context, and φ be a well-formed red proposition. We define red entailment by $\Gamma \vDash_R \varphi$ iff $\llbracket \Gamma \rrbracket \leq \llbracket \varphi \rrbracket$, i.e. iff there exists a morphism between the objects $\llbracket \Gamma \rrbracket$ and $\llbracket \varphi \rrbracket$ in \mathfrak{T}_{\otimes} .

5 Metatheorems

With a semantics for both the blue and red fragments, we are ready to take on the fundamental metatheorems of soundness and completeness. The hierarchical structure of the proof calculi (and their semantics) gives a natural separation of work, as the soundness and completeness of the red fragment depends directly, via the recoloring rule, on the corresponding theorems for the blue fragment.

Theorem 4 (Soundness). If $\Gamma \vdash_B \varphi$ then $\Gamma \vDash_B \varphi$; and if $\Gamma \vdash_R \varphi$ then $\Gamma \vDash_R \varphi$.

Both parts follow straightforwardly by induction; the only interesting case is recoloring, which follows by Lemma 2 and soundness of the blue fragment. The completeness theorems require a little more work; blue completeness relies on the Lindenbaum-Tarski method (*i.e.*, by taking the set of blue propositions quotiented by blue provable equivalence, $\Phi_B/\dashv\vdash_B$, and showing that this is isomorphic to \mathfrak{T}), while red completeness uses the characterization of objects of \mathfrak{T}_{\otimes} given by Theorem 3 as an induction principle for objects of \mathfrak{T}_{\otimes} .

Theorem 5 (Completeness). If $\Gamma \vDash_B \varphi$ then $\Gamma \vDash_B \varphi$; and if $\Gamma \vDash_R \varphi$ then $\Gamma \vDash_R \varphi$.

We are finally ready to tackle our previous obligation to show our propositional semantics equivalent to Toffoli's Boolean ring semantics. The first step is to show that Boolean rings are equivalent to Toffoli lattices:

Theorem 6 (Universality). A is a Toffoli lattice iff it is a Boolean ring.

Proof (Proof sketch). If $\mathfrak A$ is a Toffoli lattice, we define the constants and operations of a ring by

$$0 = \bot$$
 $1 = \top$ $a \cdot b = a \wedge b$ $a \oplus b = a \leftrightharpoons \overline{b}$

for all elements a and b of \mathfrak{A} . From this, it is straightforwardly shown that \mathfrak{A} forms an abelian group under addition (with each a as its own additive inverse, and 0 as unit), and a monoid under multiplication (with 1 as unit) which further distributes over addition; thus it is a ring, and that it is Boolean follows directly by the idempotence of meets.

In the other direction, suppose $\mathfrak A$ is a Boolean ring; then it is also a Boolean algebra [19], so it suffices to show that a Boolean algebra is also a Toffoli lattice. But then we can construct relative equivalences by $a \leftrightharpoons b = (\overline{a} \lor b) \lor (\overline{b} \lor a)$ for all elements $a, b \in |\mathfrak A|$; that $\mathfrak A$ is then a Toffoli lattice follows by algebraic manipulation.

We now extend this result to the full generality of entire reversible circuits. Let the order of a reversible circuit denote its number of input (equivalently output) lines; having the same order is thus a trivial requirement for two reversible circuits to be strongly equivalent, as the functions they compute (denote this function $f_{\mathcal{C}}$ for a circuit \mathcal{C}) will otherwise differ fundamentally by domain and codomain. Further, we will use $\mathfrak{B} = (\{0,1\},0,1,\oplus,\cdot)$ to denote the Boolean ring on the set $\{0,1\}$ with exclusive disjunction as addition, and conjunction as multiplication, and \mathfrak{B}^n to be the direct product of \mathfrak{B} with itself n times. Using Toffoli's Boolean ring semantics (as presented in Sec. 2, Fig. 1a), we will develop a semantic preorder on reversible circuits – but to do this, we need a way to handle ancillae (lines of constant value) in a clean way. This is done by the ancilla restriction on a circuit, defined as follows:

Definition 8. Let C be a reversible circuit of order n, and $\mathbf{x} \in |\mathfrak{B}^n|$. We define the ancilla restriction on \mathbf{x} with respect to C to be $\mathbf{x}_{|C} = (c_1, c_2, \ldots, c_n)$ where each c_i is given by

$$c_i = \begin{cases} k & \text{if the } i^{\text{th}} \text{ input of } \mathcal{C} \text{ is an ancilla of value } k \\ \pi_i(\mathbf{x}) & \text{otherwise} \end{cases}$$

This allows the following preorder on the set of reversible logic circuits, and in turn, the category induced by this preorder:

Lemma 3. The relation on reversible circuits defined by $C \leq_R \mathcal{D}$ iff $f_C(\mathbf{x}_{|C}) \leq f_{\mathcal{D}}(\mathbf{x}_{|\mathcal{D}})$ for all $\mathbf{x} \in |\mathfrak{B}^n|$ and circuits C, \mathcal{D} of equal order n, where the order relation $\cdot \leq \cdot$ denotes the usual (component-wise) ordering on Boolean vectors of length n, is a preorder.

Definition 9. Let RC denote the category which has reversible circuits as objects, and a single morphism between circuits C and D iff $C \leq_R D$.

Note particularly from this definition that objects \mathcal{C} and \mathcal{D} of \mathbf{RC} are isomorphic (i.e, $\mathcal{C} \leq_R \mathcal{D}$ and $\mathcal{D} \leq_R \mathcal{C}$) precisely when they are strongly equivalent. This allows us to show that all strong equivalences of reversible logic circuits are contained in \mathfrak{T}_{\otimes} :

Theorem 7 (Embedding of RC). There exists a functor $H : \mathbf{RC} \to \mathfrak{T}_{\otimes}$ which constitutes an embedding of \mathbf{RC} in \mathfrak{T}_{\otimes} , i.e., it is fully faithful; in particular $H(\mathcal{C}) \cong H(\mathcal{D})$ iff $\mathcal{C} \cong \mathcal{D}$.

Proof. We define $H: \mathbf{RC} \to \mathfrak{T}_{\otimes}$ on objects by taking circuits to their annotation, as given by the annotation algorithm (see Sec. 2 and the example in Fig. 2), and on morphisms by taking $\mathcal{C} \leq \mathcal{D}$ to the morphism $H(\mathcal{C}) \leq H(\mathcal{D})$: That this morphism exists in \mathfrak{T} follows by induction on the order of \mathcal{C} (equivalently \mathcal{D}) by Theorem 6, since the order on the outputs is an order on Boolean ring terms, which are equivalent to Toffoli lattice terms via

$$a \cdot b = a \wedge b \hspace{1cm} a \oplus b = a \leftrightarrows \overline{b} \hspace{1cm} a \oplus 1 = a \leftrightarrows \overline{\top} = a \leftrightarrows \bot = \overline{a}$$

which shows, by soundness, completeness and the denotation of the propositional semantics, the exact correspondence between Toffoli's Boolean ring semantics and our propositional semantics (see Sec. 2, Figs. 1a and 1b). That $H(\mathcal{C}) \cong H(\mathcal{D})$ iff $\mathcal{C} \cong \mathcal{D}$ (equivalently, that H is fully faithful) follows likewise by induction on the order of \mathcal{C} (equivalently \mathcal{D}) using Theorem 6.

6 Applications

Above, we have shown that the logic of LRS_{\circledast} is sound and complete with respect to a semantics that includes all strong equivalences of reversible logic circuits. This property suggests, as an obvious first application, a general method for proving such strong equivalences: Use the annotation algorithm of Sec. 2 to extract propositional representations of each circuit, and then use LRS_{\circledast} to show that their propositional representations are provably equivalent.

This approach can be applied directly in the optimization of reversible circuits. When used on very large circuits, the annotation algorithm may produce propositional representations that are infeasibly large to work with, however. Where the approach really shines is in the development and verification of template-based reversible circuit rewriting systems (see, e.g., [20, 1]). Template-based rewriting works by identifying certain forms of sub-circuits, allowing these to be substituted with equivalent ones.

Since such templates are typically quite modest in size, one can often extract corresponding propositions from templates with only a few steps of the annotation algorithm. A concrete example of such a template-based rewriting rule is Soeken & Thomsen's rule R2,

shown on the right. Annotating these two circuits with our algorithm, the rule states precisely the equivalences

$$\neg x_1 \circledast x_1 \bullet \neg x_2 \dashv \vdash_R \neg x_1 \circledast \neg \neg x_1 \bullet \neg x_2 \tag{1}$$

and

$$\neg x_1 \circledast \neg x_1 \bullet \neg x_2 \dashv \vdash_R \neg x_1 \circledast \neg x_1 \bullet \neg x_2 . \tag{2}$$

which are both, indeed, provable. Note that (2) follows directly by red identity, as the annotation the two circuits resulted in syntactically identical propositions. One of the two derivations proving the (1) is shown in Fig. 5.

Using diagrammatic notation for such such rewriting systems is both convenient and intuitive to use for humans. Although this has provided real insights into the rewriting behavior of reversible circuits, showing completeness (with respect to reversible circuits) for such rewriting systems has proven difficult.

Because LRS $_{\circledast}$ provides sound and complete proof calculi for reasoning about reversible circuits, we can go the other way around and extract an equational theory from this that is sound and complete with respect to reversible circuits. Further, since the blue fragment of LRS $_{\circledast}$ is sound and complete with respect to Toffoli lattices, we can instead extract an equational theory for the blue fragment from the definition of a Toffoli lattice, using the following lemma:

Fig. 5: Derivation in LRS $_{\circledast}$ for verifying the first direction of Soeken & Thomsen's rule R2.

Lemma 4. In any Toffoli lattice, $a \leq b$ iff $a \wedge \overline{b} = \bot$.

This lemma allows us to straightforwardly recast the definition of a Toffoli lattice in purely equational terms (although the result is not exactly elegant). What this *does* give us, is a set of equations that must hold for all Toffoli lattices, and which any other complete equational theory must therefore be equivalent to, and the means to show such an equivalence by converting equalities to statements about the underlying order structure and vice versa.

Fig. 6 shows a more pleasing equational theory for the blue fragment, presented in the syntax of LRS $_{\circledast}$ (the intrinsic properties of equality, *i.e.*, reflexivity, symmetry, transitivity, and congruences are not shown,) proven equivalent (and therefore sound and complete) exactly in the way outlined above (by the power of boring algebra). Deriving an equational theory for the red fragment is simpler, as it is sound and complete with respect to the free monoidal part of \mathfrak{T}_{\otimes} , which is already expressed in equational terms. As such, the equational theory for the red fragment shown in Fig. 6 is sound and complete by definition, though congruences applied in the red fragment are syntactically restricted by recolorability; that is, we can only replace recolorable propositions by recolorable propositions.

The usefulness of such an equational theory is evident in that we can, e.g., now prove the soundness of the R2 rules directly by applying equation (B9) in Fig. 6. Such equational theories can themselves also be used to develop new rewriting systems for reversible circuits, in particular to suggest new templates.

```
\varphi\circledast (\psi\circledast\chi)\overset{(\operatorname{R1})}{=}(\varphi\circledast\psi)\circledast\chi \qquad \qquad (\varphi\&\psi)\&\chi\overset{(\operatorname{B4})}{=}\varphi\&(\psi\&\chi) \varphi\circledast e\overset{(\operatorname{R2})}{=}\varphi \qquad \qquad \varphi\bullet - \psi\overset{(\operatorname{B5})}{=}\psi\bullet - \varphi e\circledast\varphi\overset{(\operatorname{R3})}{=}\varphi \qquad \qquad (\varphi\bullet - \psi)\bullet - \chi\overset{(\operatorname{B6})}{=}\varphi\bullet - (\psi\bullet - \chi) \varphi\& \neg (\psi\&\chi)\overset{(\operatorname{B7})}{=} \neg (\neg(\varphi\&\neg\psi)\&\neg(\varphi\&\neg\chi)) \varphi\& 1\overset{(\operatorname{B1})}{=}\varphi \qquad \qquad \varphi\bullet - \psi\overset{(\operatorname{B8})}{=}\neg(\varphi\&\neg\psi)\&\neg(\psi\&\neg\varphi) \neg\neg\varphi\overset{(\operatorname{B9})}{=}\varphi \varphi\&\psi\overset{(\operatorname{B3})}{=}\psi\&\varphi \qquad \qquad \varphi\&\neg\varphi\overset{(\operatorname{B10})}{=}0
```

Fig. 6: Sound and complete equational theories for the two calculi.

7 Conclusion and future work

In this article, we have presented a syntactic representation of reversible logic circuits centered around the control interpretation of Toffoli's reversible gate set, and shown, via two proof calculi of natural deduction, that a variant of classical propositional logic extended with ordered multiplicative conjunction is sufficient to reason about these. We have developed an algebraic and categorical semantics, shown that the proof calculi are sound and complete with respect to these, and that this model subsumes the established notion of strong equivalence of reversible logic circuits. Finally, we have shown how our work can be used to prove strong equivalences of reversible logic circuits, to verify existing systems of reversible logic circuit rewriting, and to develop new such rewriting systems.

The approach has been successful in enabling reasoning about reversible logic circuits, but it is not quite on even footing with the template-based approaches to reversible circuit rewriting, as these use a graphical circuit notation which, by definition, asserts circuit reversibility on every rewriting step. Although our approach faithfully models circuit semantics, it is not currently clear when looking at an arbitrary proposition whether it corresponds to a reversible circuit or not. On the other hand, by decoupling the propositions from the graphical representations, the current logic may allow for much shorter rewritings than if each step must yield representations which directly translate to circuits in this way.

References

- M. Arabzadeh, M. Saeedi, and M. S. Zamani. Rule-based optimization of reversible circuits. In Proc. ASP-DAC, 2010, pages 849–854, 2010.
- A. Bérut, A. Arakelyan, A. Petrosyan, S. Ciliberto, R. Dillenschneider, and E. Lutz. Experimental verification of Landauer's principle linking information and thermodynamics. *Nature*, 483(7388):187–189, 2012.
- 3. S. R. Buss. Handbook of Proof Theory. Elsevier, 1998.
- 4. A. De Vos. Reversible Computing. Wiley-VCH, 2010.

- A. De Vos, S. Burignat, R. Glück, T. A. Mogensen, H. B. Axelsen, M. K. Thomsen, E. Rotenberg, and T. Yokoyama. Designing garbage-free reversible implementations of the integer cosine transform. ACM J. Emerg. Tech. Com., 11(2):11:1-11:15, 2014.
- R. P. Feynman. Quantum mechanical computers. Found. Phys., 16(6):507–531, 1986.
- 7. B. Jacobs. Categorical Logic and Type Theory. Elsevier, 1999.
- 8. S. P. Jordan. Strong equivalence of reversible circuits is coNP-complete. *Quantum Inf. Comput.*, 14(15-16):1302–1307, 2014.
- A. Joyal and R. Street. The geometry of tensor calculus, i. Adv. Math., 88(1):55–112, 1991.
- 10. A. Joyal and R. Street. Braided tensor categories. Adv. Math., 102(1):20-78, 1993.
- 11. R. Kaarsgaard. Towards a propositional logic for reversible logic circuits. In *Proceedings of the ESSLLI 2014 Student Session*, pages 33-41, 2014. Available online at http://www.kr.tuwien.ac.at/drm/dehaan/stus2014/proceedings.pdf.
- 12. R. Landauer. Irreversibility and heat generation in the computing process. $IBM\ J.$ Res. Dev., 5(3):261–269, 1961.
- 13. T. Leinster. Higher operads, higher categories, volume 298 of London Mathematical Society Lecture Note Series. Cambridge University Press, 2004.
- 14. A. O. Orlov, C. S. Lent, C. C. Thorpe, G. P. Boechler, and G. L. Snider. Experimental test of Landauer's principle at the sub- $k_b t$ level. *Japan. J. Appl. Phys.*, 51:06FE10, 2012.
- 15. J. Polakow. Ordered linear logic and applications. PhD thesis, CMU, 2001.
- 16. T. Rendel and K. Ostermann. Invertible syntax descriptions: Unifying parsing and pretty printing. ACM SIGPLAN Notices, 45(11):1–12, 2010.
- 17. M. P. Schellekens. MOQA; unlocking the potential of compositional static average-case analysis. J. Log. Algebr. Program., 79(1):61–83, 2010.
- 18. R. A. G. Seely. Linear logic, *-autonomous categories and cofree coalgebras. *Contemp. Math.*, 92:371–382, 1989.
- 19. R. Sikorski. Boolean algebras. Springer, 1969.
- M. Soeken and M. K. Thomsen. White dots do matter: Rewriting reversible logic circuits. In Reversible Computation, volume 7948 of LNCS, pages 196–208, 2013.
- M. K. Thomsen, R. Glück, and H. B. Axelsen. Reversible arithmetic logic unit for quantum arithmetic. J. Phys. A: Math. Theor., 43(38):382002, 2010.
- T. Toffoli. Reversible computing. In Proc. ICALP, 1980, pages 632–644. Springer, 1980
- 23. R. Wille and R. Drechsler. Towards a design flow for reversible logic. Springer, 2010.