Logical Methods in Computer Science
Volume 16, Issue 2, 2020, pp. 8:1–8:23
https://lmcs.episciences.org/

Submitted Feb. 01, 2019
Published Jun. 02, 2020

ON PROPERTIES OF B-TERMS

MIRAI IKEBUCHI AND KEISUKE NAKANO

Massachusetts Institute of Technology, Cambridge, MA, USA
e-mail address: ikebuchi@mit.edu

Tohoku University, Sendai, Miyagi, Japan
e-mail address: k.nakano@acm.org

Abstract. B-terms are built from the B combinator alone defined by $B \equiv \lambda fgx.f(g\ x)$, which is well known as a function composition operator. This paper investigates an interesting property of B-terms, that is, whether repetitive right applications of a B-term cycles or not. We discuss conditions for B-terms to have and not to have the property through a sound and complete equational axiomatization. Specifically, we give examples of B-terms which have the cyclic property and show that there are infinitely many B-terms which do not have the property. Also, we introduce another interesting property about a canonical representation of B-terms that is useful to detect cycles, or equivalently, to prove the cyclic property, with an efficient algorithm.

Introduction

The 'bluebird' combinator $B = \lambda fgx.\ f(g\ x)$ is well known [Sch24, Cur30a, Smu12] as a bracketing combinator or composition operator, which has a principal type $(\alpha \to \beta) \to (\gamma \to \alpha) \to \gamma \to \beta$. A function $B\ f\ g$ (also written as $f \circ g$) takes a single argument x and returns the term $f(g\ x)$. In the general case that g takes n arguments, the composition can be given by $\lambda x_1 \ldots x_n.\, f(g\ x_1\ \ldots\ x_n)$. We call it the *n-argument composition* of f and g. Interestingly, the function can be given as $B^n\ f\ g$ where e^n stands for the n-fold composition $\underbrace{e \circ \cdots \circ e}_{n}$ of the function e, or equivalently defined by $e^n\, x = \underbrace{e\ (\ldots (e}_{n}\ x))$. This fact can be shown by an easy induction.

Now we consider the 2-argument composition expressed as $B^2 = \lambda fgxy.\, f(g\ x\ y)$. From the definition, we have $B^2 = B \circ B = B\ B\ B$. Note that function application is considered left-associative, that is, $f\ a\ b = (f\ a)\ b$. Thus B^2 is expressed as a term in which all applications nest to the left, never to the right. We call such terms *flat* [Oka03]. We write $X_{(k)}$ for the flat term defined by $\underbrace{X\ X\ X\ \ldots\ X}_{k} = \underbrace{(\ldots((X\ X)\ X)\ldots)\ X}_{k}$ (that can be written as $I\, X^{\sim k}$ in Barendregt's notation [Bar84]). Using this notation, we can write $B^2 = B_{(3)}$.

Key words and phrases: Combinatory logic, B combinator, Lambda calculus.

LOGICAL METHODS IN COMPUTER SCIENCE
DOI:10.23638/LMCS-16(2:8)2020

$$\begin{tikzcd}[arrows=dash]
B_{(1)} \arrow[r] & B_{(2)} \arrow[r] & B_{(3)} \arrow[r] & B_{(4)} \arrow[r] & B_{(5)} \arrow[r] & B_{(6)} \; (= B_{(10)} = B_{(14)} = \dots) \arrow[r] \arrow[d] & B_{(7)} \; (= B_{(11)} = B_{(15)} = \dots) \arrow[d] \\
 & & & & & B_{(9)} \; (= B_{(13)} = B_{(17)} = \dots) \arrow[r] & B_{(8)} \; (= B_{(12)} = B_{(16)} = \dots)
\end{tikzcd}$$

FIGURE 1. ρ-property of the B combinator

Okasaki [Oka03] investigated facts about flatness. For example, he shows that there is no universal combinator X that can represent any combinator by $X_{(k)}$ with some k. We shall delve into the case of $X = B$. Consider the n-argument composition operator B^n. We have already seen that B^2 is $\beta\eta$-equivalent to the flat term $B_{(3)}$. For $n = 3$, using $\underline{f}\ (\underline{\underline{g}}\ x) = B\ f\ g\ x$, we have

$$\begin{aligned}
B^3 &= B\ B^2\ B \\
&= \underline{B}\ (\underline{\underline{B\ B}}\ B)\ B \\
&= \underline{B\ B}\ (\underline{\underline{B}}\ B)\ B\ B \\
&= \underline{B}\ (\underline{\underline{B}}\ B)\ B\ B\ B\ B \\
&= B\ B\ B\ B\ B\ B\ B\ B,
\end{aligned}$$

and thus $B^3 = B_{(8)}$. How about the 4-argument composition B^4? In fact, there is no integer k such that $B^4 = B_{(k)}$ with respect to $\beta\eta$-equality. Moreover, for any $n > 3$, there does not exist k such that $B^n = B_{(k)}$. This surprising fact is proved by a quite simple method; listing all $B_{(k)}$s for $k = 1, 2, \dots$ and checking that none of them is equivalent to B^n. An easy computation gives $B_{(6)} = B_{(10)} = \lambda xyzwv.\ x\ (y\ z)\ (w\ v)$, and hence $B_{(i)} = B_{(i+4)}$ for every $i \geq 6$. Then, by computing $B_{(k)}$s only for $k \in \{1, 2, \dots, 6\}$, we can check that $B_{(k)}$ is not $\beta\eta$-equivalent to B^n with $n > 3$ for $k \in \{1, 2, \dots\}$. Thus we conclude that there is no integer k such that $B^n = B_{(k)}$.

This is the starting point of our research. We say that a combinator X has the *ρ-property* if there exist two distinct integers i and j such that $X_{(i)} = X_{(j)}$. In this case, we have $X_{(i+k)} = X_{(j+k)}$ for any $k \geq 0$ (à la *finite monogenic semigroup* [Lja68]). Fig. 1 shows a computation graph of $B_{(k)}$. The ρ-property is named after the shape of the graph.

This paper discusses the ρ-property of combinatory terms, particularly terms built from B alone. We call such terms *B-terms* and $\mathbf{CL}(B)$ denotes the set of all B-terms. For example, the B-term $B\ B$ enjoys the ρ-property with $(B\ B)_{(52)} = (B\ B)_{(32)}$ and so does $B\ (B\ B)$ with $(B\ (B\ B))_{(294)} = (B\ (B\ B))_{(258)}$ as reported in [Nak08]. Several combinators other than B-terms can be found to enjoy the ρ-property, for example, $K = \lambda xy.\ x$ and $C = \lambda xyz.\ x\ z\ y$ because of $K_{(3)} = K_{(1)}$ and $C_{(4)} = C_{(3)}$. They are less interesting in the sense that the cycle starts immediately and its size is very small, comparing with B-terms like $B\ B$ and $B\ (B\ B)$. As we will see later, $B\ (B\ (B\ (B\ (B\ (B\ B)))))(\equiv B^6\ B)$ has the ρ-property with the cycle of size more than 3×10^{11} which starts after more than 2×10^{12} repetitive right applications. This is why the ρ-property of B-terms is intensively discussed in the present paper. A general definition of the ρ-property is presented in Section 1.

The contributions of the paper are two-fold. One is to give a characterization of $\mathbf{CL}(B)$ (Section 2) and another is to provide a sufficient condition for the ρ-property and anti-ρ-property of B-terms (Section 3). In the former, we introduce a canonical representation of B-terms and establish a sound and complete equational axiomatization for $\mathbf{CL}(B)$. In the latter, the ρ-property of $B^n B$ with $n \leq 6$ is shown with an efficient algorithm and the anti-ρ-property for B-terms of particular forms is proved.

This paper extends and refines our paper presented in FSCD 2018 [IN18]. Compared to our previous work, we have made several improvements. First, we add relationships to the existing work, the Curry's compositive normal form and the Thompson's group. Second, we report progress on proving and disproving the ρ-property of B-terms. For proving the ρ-property, we add more precise information on the implementation of our ρ-property checker. For disproving the ρ-property, we introduce another proof method for a specific B-term and expand the set of B-terms which are known not to have the ρ-property. Furthermore, we discuss other possible approaches for further steps to show a conjecture by the second author [Nak08].

1. The ρ-property of terms

The ρ-property of a combinator X is that $X_{(i)} = X_{(i+j)}$ holds for some $i, j \geq 1$. We adopt $\beta\eta$-equality of corresponding λ-terms for the equality of combinatory terms in this paper. We could use another equality, for example, induced by the axioms of combinatory logic. The choice of equality is not essential here, e.g., $B_{(9)}$ and $B_{(13)}$ are equal even up to the combinatory axiom of B, as well as $\beta\eta$-equality. (See Section 4 for more details.) Furthermore, for simplicity, we only deal with the case where $X_{(n)}$ is normalizable for all n. If $X_{(n)}$ is not normalizable, it is much more difficult to check equivalence with the other terms. This restriction does not affect the results of the paper because all B-terms are normalizing.

Let us write $\rho(X) = (i, j)$ if a combinator X has the ρ-property due to $X_{(i)} = X_{(i+j)}$ with minimum positive integers i and j. For example, we have $\rho(B) = (6, 4)$, $\rho(C) = (3, 1)$, $\rho(K) = (1, 2)$ and $\rho(I) = (1, 1)$. Besides them, several combinators introduced in Smullyan's book [Smu12] have the ρ-property:

$$\begin{aligned}
\rho(D) &= (32, 20) && \text{where } D = \lambda xyzw.\, x\ y\ (z\ w) \\
\rho(F) &= (3, 1) && \text{where } F = \lambda xyz.\, z\ y\ x \\
\rho(R) &= (3, 1) && \text{where } R = \lambda xyz.\, y\ z\ x \\
\rho(T) &= (2, 1) && \text{where } T = \lambda xy.\, y\ x \\
\rho(V) &= (3, 1) && \text{where } V = \lambda xyz.\, z\ x\ y.
\end{aligned}$$

Except for the B and D $(= B\ B)$ combinators, the property is 'trivial' in the sense that the loop starts early and the size of the cycle is very small.

On the other hand, the combinators $S = \lambda xyz.\, x\ z\ (y\ z)$ and $O = \lambda xy.\, y\ (x\ y)$ in the book do not have the ρ-property since their right application expands the λ-terms as illustrated by

$$\begin{aligned}
S_{(2n+1)} &= \lambda xy.\, \underbrace{x\ y\ (x\ y\ (\ldots (x\ y}_{n}\ (\lambda z.x\ z\ (y\ z))) \ldots)), \\
O_{(n+1)} &= \lambda x.\, \underbrace{x\ (x\ (\ldots (x}_{n}\ (\lambda y.y\ (x\ y)).
\end{aligned}$$

The definition of the ρ-property is naturally extended from single combinators to terms obtained by combining several combinators. We found by computation that several B-terms have the ρ-property as shown below.

$$\rho(B^0B) = (6,4) \qquad \rho(B^4B) = (191206, 431453)$$
$$\rho(B^1B) = (32,20) \qquad \rho(B^5B) = (766241307, 234444571)$$
$$\rho(B^2B) = (258,36) \qquad \rho(B^6B) = (2641033883877, 339020201163)$$
$$\rho(B^3B) = (4240,5796)$$

The details will be shown in Section 3.1.

From his observation on repetitive right applications for several B-terms, Nakano [Nak08] has conjectured as follows.

Conjecture 1.1. A B-term e has the ρ-property if and only if e is a monomial, i.e., e is equivalent to B^nB with $n \geq 0$.

The definition of *monomial* will be given in Section 2.3 in the context of a canonical representation of B-terms. The "if" part of the conjecture for $n \leq 6$ is shown by the above results; the "only if" part will be shown for specific B terms which will be discussed Section 3.2. Note that the ρ-property of X can be rephrased in terms of the set generated by right application, that is, the finiteness of the set $\{nf(X_{(n)}) \mid n \geq 1\}$ where $nf(e)$ represents the normal form of e. Conjecture 1.1 claims that for any B-term e, the finiteness of the set $\{nf(e_{(n)}) \mid n \geq 1\}$ is decidable since so is the word problem of B-terms.

2. Checking equivalence of B-terms

The set of all B-terms, $\mathbf{CL}(B)$, is closed under application by definition, that is, the repetitive right application of a B-term always generates a sequence of B-terms. Hence, the ρ-property can be decided by checking 'equivalence' among generated B-terms, where the equivalence should be checked through $\beta\eta$-equivalence of their corresponding λ-terms in accordance with the definition of the ρ-property. It would be useful if we have a fast algorithm for deciding equivalence over B-terms.

In this section, we give a characterization of the B-terms to efficiently decide their equivalence. We introduce a method for deciding equivalence of B-terms without calculating the corresponding λ-terms. To this end, we first investigate equivalence over B-terms with examples and then present an equation system as a characterization of B-terms so as to decide equivalence between two B-terms. Based on the equation system, we introduce a canonical representation of B-terms. The representation makes it easy to observe the growth caused by repetitive right application of B-terms, which will be later used for proving the anti-ρ-property of B^2. We believe that this representation will be helpful to prove the ρ-property or the anti-ρ-property for the other B-terms.

2.1. Equivalence over B-terms.

Two B-terms are said to be *equivalent* if their corresponding λ-terms are $\beta\eta$-equivalent. For instance, $B\ B\ (B\ B)$ and $B\ (B\ B)\ B\ B$ are equivalent. This can be shown by the definition $B\ x\ y\ z = x\ (y\ z)$. For another (non-trivial) instance, $B\ B\ (B\ B)$ and $B\ (B\ (B\ B))\ B$ are equivalent. This is illustrated by the fact that they are equivalent to $\lambda xyzwv.\, x\ (y\ z)\ (w\ v)$ where B is replaced with $\lambda xyz.\, x\ (y\ z)$ or the other way around at the $=_\beta$ equation. Similarly, it is hard to directly show equivalence between the two B-terms, $B\ (B\ B)\ (B\ B)$ and $B\ (B\ B\ B)$, which requires long calculation like:

$$\begin{aligned}
B\ B\ (B\ B) &=_\eta \lambda x.B\ B\ (B\ B)\ x \\
&=_\beta \lambda x.B\ (B\ B\ x) \\
&=_\eta \lambda xy.B\ (B\ B\ x)\ y \\
&=_\eta \lambda xyz.B\ (B\ B\ x)\ y\ z \\
&=_\beta \lambda xyz.B\ B\ x\ (y\ z) \\
&=_\beta \lambda xyz.B\ (x\ (y\ z)) \\
&=_\beta \lambda xyz.B\ (B\ x\ y\ z) \\
&=_\beta \lambda xyz.B\ B\ (B\ x\ y)\ z \\
&=_\eta \lambda xy.B\ B\ (B\ x\ y) \\
&=_\beta \lambda xy.B\ (B\ B)\ (B\ x)\ y \\
&=_\eta \lambda x.B\ (B\ B)\ (B\ x) \\
&=_\beta \lambda x.B\ (B\ (B\ B))\ B\ x \\
&=_\eta B\ (B\ (B\ B))\ B.
\end{aligned}$$

This kind of equality makes it hard to investigate the ρ-property of B-terms. To solve this annoying issue, we will introduce a canonical representation of B-terms in Section 2.3.

2.2. **Equational axiomatization for B-terms.** Equality between two B-terms can be decided through their canonical representation introduced in Section 2.3. The representation is based on a sound and complete equation system as described in the next theorem.

Theorem 2.1. *Two B-terms are $\beta\eta$-equivalent if and only if their equality is derived from the following equations:*

$$B\ x\ y\ z = x\ (y\ z) \tag{B1}$$

$$B\ (B\ x\ y) = B\ (B\ x)\ (B\ y) \tag{B2}$$

$$B\ B\ (B\ x) = B\ (B\ (B\ x))\ B \tag{B3}$$

The proof of the "if" part, which corresponds to the soundness of the equation system (B1), (B2), and (B3), is given here. We will later prove the "only if" part with the uniqueness of the canonical representation of B-terms.

Proof. Equation (B1) is immediate from the definition of B. Equations (B2) and (B3) are shown by

$$\begin{aligned} B\ (B\ e_1\ e_2) &= \lambda xy.\, B\ (B\ e_1\ e_2)\ x\ y \\ &= \lambda xy.\, B\ e_1\ e_2\ (x\ y) \\ &= \lambda xy.\, e_1\ (e_2\ (x\ y)) \\ &= \lambda xy.\, e_1\ (B\ e_2\ x\ y) \\ &= \lambda x.\, B\ e_1\ (B\ e_2\ x) \\ &= B\ (B\ e_1)\ (B\ e_2) \end{aligned} \qquad \begin{aligned} B\ B\ (B\ e_1) &= \lambda x.\, B\ B\ (B\ e_1)\ x \\ &= \lambda x.\, B\ (B\ e_1\ x) \\ &= \lambda xyz.\, B\ e_1\ x (y\ z) \\ &= \lambda xyz.\, e_1\ (x\ (y\ z)) \\ &= \lambda xyz.\, e_1\ (B\ x\ y\ z) \\ &= \lambda xy.\, B\ e_1\ (B\ x\ y) \\ &= \lambda x.\, B\ (B\ e_1)\ (B\ x) \\ &= B\ (B\ (B\ e_1))\ B \end{aligned}$$

where the α-renaming is implicitly used. □

Equation (B2) has been employed by Statman [Sta11] to show that no $B\omega$-term can be a fixed-point combinator where $\omega = \lambda x.\, x\ x$. This equation exposes an interesting feature of the B combinator. Write equation (B2) as

$$B\ (e_1 \circ e_2) = (B\ e_1) \circ (B\ e_2) \tag{B2'}$$

by replacing every B combinator with $\circ$ infix operator if it has exactly two arguments. The equation is a distributive law of B over $\circ$, which will be used to obtain the canonical representation of B-terms. Equation (B3) is also used for the same purpose as the form of

$$B \circ (B\ e_1) = (B\ (B\ e_1)) \circ B. \tag{B3'}$$

We also have a natural equation $B\ e_1\ (B\ e_2\ e_3) = B\ (B\ e_1\ e_2)\ e_3$ which represents associativity of function composition, i.e., $e_1 \circ (e_2 \circ e_3) = (e_1 \circ e_2) \circ e_3$. This is shown with equations (B1) and (B2) by

$$B\ e_1\ (B\ e_2\ e_3) = B\ (B\ e_1)\ (B\ e_2)\ e_3 = B\ (B\ e_1\ e_2)\ e_3.$$

2.3. **Canonical representation of B-terms.** To decide equality between two B-terms, it does not suffice to compute their normal forms under the definition of B, $B\ x\ y\ z \to x\ (y\ z)$. This is because two distinct normal forms may be equal up to $\beta\eta$-equivalence, e.g., $B\ B\ (B\ B)$ and $B\ (B\ (B\ B))\ B$. We introduce a canonical representation of B-terms, which makes it easy to check equivalence of B-terms. We will eventually find that for any B-term e there exists a unique finite non-empty weakly-decreasing sequence of non-negative integers $n_1 \geq n_2 \geq \cdots \geq n_k$ such that e is equivalent to $(B^{n_1} B) \circ (B^{n_2} B) \circ \cdots \circ (B^{n_k} B)$. Ignoring the inequality condition gives *polynomials* introduced by Statman [Sta11]. We will use these *decreasing polynomials* for our canonical representation as presented later. A similar result is found in [Cur30b] as discussed later.

First, we explain how this canonical form is obtained from a B-term. We only need to consider B-terms in which every B has at most two arguments. One can reduce the arguments of B to less than three by repeatedly rewriting occurrences of $B\ e_1\ e_2\ e_3\ e_4\ \ldots\ e_n$ into $e_1\ (e_2\ e_3)\ e_4\ \ldots\ e_n$. The rewriting procedure always terminates because it reduces the number of B. Thus, every B-term in $\mathbf{CL}(B)$ is equivalent to a B-term built by the syntax

$$e ::= B \ \mid\ B\ e \ \mid\ e\ \circ\ e \tag{2.1}$$

where $e_1 \circ e_2$ denotes $B\ e_1\ e_2$. We prefer to use the infix operator $\circ$ instead of B that has two arguments because associativity of B, that is, $B\ e_1\ (B\ e_2\ e_3) = B\ (B\ e_1\ e_2)\ e_3$ can be implicitly assumed. This simplifies the further discussion on B-terms. We will deal with only B-terms in syntax (2.1) from now on. The $\circ$ operator has lower precedence than application in this paper, e.g., terms $B\ B \circ B$ and $B \circ B\ B$ represent $(B\ B) \circ B$ and $B \circ (B\ B)$, respectively.

The syntactic restriction by (2.1) does not suffice to proffer a canonical representation of B-terms. For example, both of the two B-terms $B \circ B\ B$ and $B\ (B\ B) \circ B$ are given in the form of (2.1), but we can see that they are equivalent using (B3').

A *polynomial form of B-terms* is obtained by putting a restriction on the syntax so that no B combinator occurs outside of the $\circ$ operator while syntax (2.1) allows the B combinators and the $\circ$ operators to occur in an arbitrary position. The restricted syntax is given as

$$e ::= e_B \ \mid\ e \circ e \qquad e_B ::= B\ \mid B\ e_B$$

where terms in e_B have a form of $B(\dots(B(B\ B))\dots)$, that is $B^n B$ with some n, called *monomial*. The syntax can be simply rewritten into $e ::= B^n B \ \mid\ e \circ e$, which is called *polynomial*.

Definition 2.2. A B-term $B^n B$ is called *monomial*. A *polynomial* is a B-term given in the form of

$$(B^{n_1} B) \circ (B^{n_2} B) \circ \cdots \circ (B^{n_k} B)$$

where $k > 0$ and $n_1, \dots, n_k \geq 0$ are integers. In particular, a polynomial is called *decreasing* when $n_1 \geq n_2 \geq \cdots \geq n_k$. The *length* of a polynomial P is the number of monomials in P, i.e., the length of the polynomial above is k. The numbers $n_1, n_2, \dots, n_k$ are called *degrees*.

In the rest of this subsection, we prove that for any B-term e there exists a unique decreasing polynomial equivalent to e. First, we show that e has an equivalent polynomial.

Lemma 2.3 [Sta11]**.** *For any B-term e, there exists a polynomial equivalent to e.*

Proof. We prove the statement by induction on the structure of e. In the case of $e \equiv B$, the term itself is polynomial. In the case of $e \equiv B\ e_1$, assume that e_1 has equivalent polynomial $(B^{n_1} B) \circ (B^{n_2} B) \circ \cdots \circ (B^{n_k} B)$. Repeatedly applying equation (B2') to $B\ e_1$, we obtain a polynomial equivalent to $B\ e_1$ as $(B^{n_1+1} B) \circ (B^{n_2+1} B) \circ \cdots \circ (B^{n_k+1} B)$. In the case of $e \equiv e_1 \circ e_2$, assume that e_1 and e_2 have equivalent polynomials P_1 and P_2, respectively. A polynomial equivalent to e is given by $P_1 \circ P_2$. □

Next, we show that for any polynomial P there exists a decreasing polynomial equivalent to P. A key equation of the proof is

$$(B^m B) \circ (B^n B) = (B^{n+1} B) \circ (B^m B) \quad \text{when } m < n, \tag{2.2}$$

which is shown by

$$\begin{aligned}(B^m B) \circ (B^n B) &= B^m (B \circ (B^{n-m} B)) \\ &= B^m (B \circ (B\ (B^{n-m-1} B))) \\ &= B^m ((B(B(B^{n-m-1} B))) \circ B) \\ &= (B^{n+1} B) \circ (B^m B)\end{aligned}$$

using equations (B2') and (B3').

Lemma 2.4. *Any polynomial P has an equivalent decreasing polynomial P' such that*
- *the length of P and P' are equal, and*
- *the lowest degrees of P and P' are equal.*

Proof. We prove the statement by induction on the length of P. When the length is 1, that is, P is a monomial, P itself is decreasing and the statement holds. When the length k of P is greater than 1, take P_1 such that $P \equiv P_1 \circ (B^n B)$. From the induction hypothesis, there exists a decreasing polynomial $P_1' \equiv (B^{n_1} B) \circ (B^{n_2} B) \circ \cdots \circ (B^{n_{k-1}} B)$ equivalent to P_1, and the lowest degree of P_1 is n_{k-1}. If $n_{k-1} \geq n$, then $P' \equiv P_1' \circ (B^n\ B)$ is decreasing and equivalent to P. Since the lowest degrees of P and P' are n, the statement holds. If $n_{k-1} < n$, P is equivalent to

$$(B^{n_1}\ B) \circ \cdots \circ (B^{n_{k-1}} B) \circ (B^n B) = (B^{n_1} B) \circ \cdots \circ (B^{n+1} B) \circ (B^{n_{k-1}} B)$$

due to equation (2.2). Putting the last term as $P_2 \circ (B^{n_{k-1}} B)$, the length of P_2 is $k-1$ and the lowest degree of P_2 is greater than or equal to n_{k-1}. From the induction hypothesis, P_2 has an equivalent decreasing polynomial P_2' of length $k-1$ and the lowest degree of P_2' greater than or equal to n_{k-1}. Thereby we obtain a decreasing polynomial $P_2' \circ (B^{n_{k-1}} B)$ equivalent to P and the statement holds. □

Example 2.5. Consider a B-term $e = B\ (B\ B\ B)\ (B\ B)\ B$. First, applying equation (B1),

$$e = B\ (B\ B\ B)\ (B\ B)\ (B\ B) = B\ B\ B\ (B\ B\ (B\ B)) = B\ (B\ (B\ B\ (B\ B)))$$

so that every B has at most two arguments. Then replacing each two-argument B to the infix $\circ$ operator, obtain $B\ (B\ (B \circ (B\ B)))$. Applying equation (B2'), we have

$$\begin{aligned} B\ (B\ (B \circ (B\ B))) &= B\ ((B\ B) \circ (B\ (B\ B))) \\ &= (B\ (B\ B)) \circ (B\ (B\ (B\ B))) \\ &= (B^2 B) \circ (B^3 B). \end{aligned}$$

Applying equation (2.2), we obtain the decreasing polynomial $(B^4 B) \circ (B^2 B)$ equivalent to e.

Every B-term has at least one equivalent decreasing polynomial as shown so far. To conclude this subsection, we show the uniqueness of decreasing polynomial equivalent to any B-term, that is, every B-term e has no two distinct decreasing polynomials equivalent to e.

The proof is based on the idea that B-terms correspond to unlabeled binary trees. Let M be a term which is constructed from variables $x_1, \ldots, x_k$ and their applications. Then we can show that if the λ-term $\lambda x_1 \ldots x_k.\, M$ is in $\mathbf{CL}(B)$, then M is obtained by putting parentheses to some positions in the sequence $x_1\ \ldots\ x_k$. More precisely, we have the following lemma.

Lemma 2.6. *Every B-term is $\beta\eta$-equivalent to a λ-term of the form $\lambda x_1 \ldots x_k.\, M$ with some $k > 2$ where M satisfies the following two conditions: (1) M consists of only the variables $x_1, \ldots, x_k$ and their applications, and (2) for every subterm of M which is in the form of $M_1\ M_2$, if M_1 has a variable x_i, then M_2 does not have any variable x_j with $j \leq i$.*

Proof. We prove the statement by induction. In the case of $e \equiv B$, e is equivalent to $\lambda x_1 x_2 x_3.\, x_1(x_2 x_3)$, the statement holds. In the case of $e \equiv B\ e_1$, e is equivalent to $\lambda x_1 x_2.\, e_1\ (x_1\ x_2)$. From the induction hypothesis, e_1 is equivalent to $\lambda x_1 x_2 \ldots x_k.\, e_1'$ where e_1' satisfies the conditions (1) and (2). Then, we can see that $e_1'\ [(x_1\ x_2)/x_1]$ also satisfies the conditions (1) and (3). □

From this lemma, we see that we do not need to specify variables in M and we can simply write like $\star\ \star\ (\star\ \star) =\ x_1\ x_2\ (x_3\ x_4)$. Formally speaking, every λ-term in $\mathbf{CL}(B)$ uniquely corresponds to a term built from $\star$ alone by the map $(\lambda x_1 \ldots x_k.\, M) \mapsto M[\star/x_1, \ldots, \star/x_k]$. We say an unlabeled binary tree (or simply, binary tree) for a term built from $\star$ alone since every term built from $\star$ alone can be seen as an unlabeled binary tree. (A term $\star$ corresponds to a leaf and $t_1\ t_2$ corresponds to the tree with left subtree t_1 and right subtree t_2.) To specify the applications in binary trees, we write $\langle t_1, t_2\rangle$ for the application $t_1\ t_2$. For example, B-terms $B = \lambda xyz.\, x\ (y\ z)$ and $B\ B = \lambda xyzw.\, x\ y\ (z\ w)$ are represented by $\langle\star, \langle\star, \star\rangle\rangle$ and $\langle\langle\star, \star\rangle, \langle\star, \star\rangle\rangle$, respectively.

We will present an algorithm for constructing the corresponding decreasing polynomial from a given binary tree. First let us define a function $\mathcal{L}_i$ with integer i which maps binary trees to lists of integers:

$$\mathcal{L}_i(\star) = [\ ] \qquad\qquad \mathcal{L}_i(\langle t_1, t_2\rangle) = \mathcal{L}_{i+\|t_1\|}(t_2) +\!\!+ \mathcal{L}_i(t_1) +\!\!+ [i]$$

where $+\!\!+$ concatenates two lists and $\|t\|$ denotes the number of leaves. For example, $\mathcal{L}_0(\langle\langle\star, \star\rangle, \langle\star, \star\rangle\rangle) = [2, 0, 0]$ and $\mathcal{L}_1(\langle\langle\star, \langle\star, \star\rangle\rangle, \langle\star, \langle\star, \star\rangle\rangle\rangle) = [4, 4, 2, 1, 1]$. Informally, the $\mathcal{L}_i$ function returns a list of integers which is obtained by labeling both leaves and nodes in the following steps. First each leaf of a given tree is labeled by $i, i+1, i+2, \ldots$ in left-to-right order. Then each internal node of the tree is labeled by the same label as its leftmost descendant leaf. The $\mathcal{L}_i$ functions return a list of labels of internal nodes in decreasing order. Figure 2 shows three examples of labeled binary trees obtained by this labeling procedure for $i = -1$. Let t_j $(j = 1, 2, 3)$ be the unlabeled binary tree corresponding to e_j. From the labeled binary trees in Figure 2, we have $\mathcal{L}_{-1}(t_1) = [1, -1, -1]$, $\mathcal{L}_{-1}(t_2) = [3, 1, 1, -1, -1]$, and $\mathcal{L}_{-1}(t_3) = [5, 2, 2, 2, 0, -1, -1, -1]$. One may notice that a binary tree t_3' corresponding η-equivalent terms of e_3 is obtained by removing the leaf 7 and its root. From $\mathcal{L}_{-1}(t_3') +\!\!+ [-1] = \mathcal{L}_{-1}(t_3)$, we have $\mathcal{L}(t_3') = \mathcal{L}(t_3)$. It is easy to show that the $\mathcal{L}$ function returns the same values for η-equivalent B-terms. The length of the list equals the number of nodes, that is, smaller by one than the number of variables in the λ-term.

Definition 2.7. $\mathcal{L}$ is the function which takes a binary tree t and returns the list of non-negative integers in $\mathcal{L}_{-1}(t)$, that is, the list obtained by excluding trailing all -1's in $\mathcal{L}_{-1}(t)$.

The following lemma claims that the $\mathcal{L}$ function computes a list of degrees of a decreasing polynomial corresponding to a given λ-term.

Lemma 2.8. *A decreasing polynomial $(B^{n_1}B) \circ (B^{n_2}B) \circ \cdots \circ (B^{n_k}B)$ is $\beta\eta$-equivalent to the λ-term $e \in \mathbf{CL}(B)$ corresponding to a binary tree t such that $\mathcal{L}(t) = [n_1, n_2, \ldots, n_k]$.*

Proof. We prove the statement by induction on the length of the polynomial P.

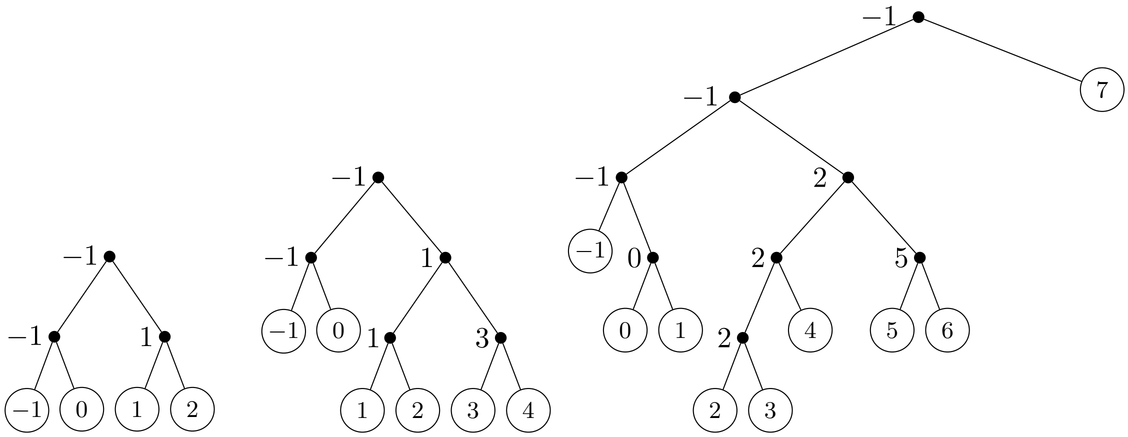
When $P \equiv B^n B$ with $n \geq 0$, it is found to be equivalent to the λ-term

$$\lambda x_1 x_2 x_3 \ldots x_{n+1} x_{n+2} x_{n+3}.\, x_1\ x_2\ x_3\ \ldots\ x_{n+1}\ (x_{n+2}\ x_{n+3})$$

by induction on n. This λ-term corresponds to the binary tree $t = \langle\langle \ldots \langle\langle\star, \underbrace{\star\rangle, \star\rangle, \ldots, \star\rangle}_{n \text{ leaves}}, \langle\star, \star\rangle\rangle$.

Then we have that $\mathcal{L}(t) = [n]$ holds from $\mathcal{L}_{-1}(t) = [n, \underbrace{-1, -1, \ldots, -1}_{n+1}]$.

When $P \equiv P' \circ (B^n B)$ with $P' \equiv (B^{n_1}B) \circ \cdots \circ (B^{n_k}B)$, $k \geq 1$ and $n_1 \geq \cdots \geq n_k \geq n \geq 0$, there exists a λ-term $\beta\eta$-equivalent to P' corresponding to a binary tree t' such that



(a) Binary tree t_1 for λ-term e_1 (b) Binary tree t_2 for λ-term e_2 (c) Binary tree t_3 for λ-term e_3

$$\begin{aligned}
\text{where}\quad & e_1 = \lambda x_1 x_2 x_3 x_4.\, x_1\ x_2\ (x_3\ x_4) \\
& e_2 = \lambda x_1 x_2 x_3 x_4 x_5 x_6.\, x_1\ x_2\ (x_3\ x_4\ (x_5\ x_6)) \\
& e_3 = \lambda x_1 x_2 x_3 x_4 x_5 x_6 x_7 x_8 x_9.\, x_1\ (x_2\ x_3)\ (x_4\ x_5\ x_6\ (x_7\ x_8))\ x_9 \\
& \quad (=_\eta \lambda x_1 x_2 x_3 x_4 x_5 x_6 x_7 x_8.\, x_1\ (x_2\ x_3)\ (x_4\ x_5\ x_6\ (x_7\ x_8))) \\
\text{and}\quad & \mathcal{L}(t_1) = [1] \\
& \mathcal{L}(t_2) = [3,1,1] \\
& \mathcal{L}(t_3) = [5,2,2,2,0]
\end{aligned}$$

Figure 2. Labeled binary trees

$\mathcal{L}(t') = [n_1, \ldots, n_k]$ from the induction hypothesis. The binary tree t' must have the form of $\langle\langle\langle \ldots \langle \underbrace{\langle \star, \star\rangle, \star\rangle, \ldots, \star\rangle}_{n_k \text{ leaves}}, t_1\rangle, \ldots, t_m\rangle$ with $m \geq 1$ and some trees $t_1, \ldots, t_m$, otherwise $\mathcal{L}(t')$ would contain an integer smaller than n_k. From the definition of $\mathcal{L}$ and $\mathcal{L}_i$, we have

$$\mathcal{L}(t') = \mathcal{L}_{s_m}(t_m) +\!\!+ \cdots +\!\!+ \mathcal{L}_{s_1}(t_1) \tag{2.3}$$

where $s_j = n_k + 1 + \sum_{i=1}^{j-1} \|t_i\|$. Additionally, the structure of t' implies $P' = \lambda x_1 \ldots x_l.\ x_1\ x_2\ \ldots\ x_{n_k+1}\ e_1 \ldots e_m$ where e_i corresponds to a binary tree t_i for $i = 1, \ldots, m$. From $B^n\ B = \lambda y_1 \ldots y_{n+3}.\, y_1\ y_2 \ldots y_{n+1}\ (y_{n+2}\ y_{n+3})$, we compute a λ-term $\beta\eta$-equivalent to $P \equiv P' \circ (B^n B)$ by

$$\begin{aligned}
P &= \lambda x.\ P'(B^n B\ x) \\
&= \lambda x.\ (\lambda x_1 \ldots x_l.\ x_1\ x_2 \ldots x_{n_k+1}\ e_1\ \ldots\ e_m) \\
&\qquad\quad (\lambda y_2 \ldots y_{n+3}.\ x\ y_2 \ldots y_{n+1}\ (y_{n+2}\ y_{n+3})) \\
&= \lambda x x_2 \ldots x_l.\ (\lambda y_2 \ldots y_{n+3}.\ x\ y_2 \ldots y_{n+1}\ (y_{n+2}\ y_{n+3}))\ x_2 \ldots x_{n_k+1}\ e_1 \ldots e_m \\
&= \lambda x x_2 \ldots x_l. \\
&\qquad (\lambda y_{n+1} y_{n+2} y_{n+3}.\ x\ x_2 \ldots x_n\ y_{n+1}\ (y_{n+2}\ y_{n+3}))\ x_{n+1} \ldots x_{n_k+1}\ e_1 \ldots e_m
\end{aligned}$$

where $n_k \geq n$ is taken into account. We split the proof into four cases: (i) $n_k = n$ and $m = 1$, (ii) $n_k = n$ and $m > 1$, (iii) $n_k = n + 1$, and (iv) $n_k > n + 1$. In the case (i) where $n_k = n$ and $m = 1$, we have

$$P = \lambda x x_2 \ldots x_l y_{n+3}.\ x\ x_2 \ldots x_n\ x_{n+1}\ (e_1\ y_{n+3}).$$

whose corresponding binary tree t is $\langle\langle \ldots \langle\langle \star, \underbrace{\star\rangle, \star\rangle, \ldots, \star\rangle}_{n \text{ leaves}}, \langle t_1, \star\rangle\rangle$. From equation (2.3), $\mathcal{L}(t) = \mathcal{L}_{n+1}(t_1) +\!\!+ [n+1] = \mathcal{L}(t') +\!\!+ [n+1] = [n_1, \ldots, n_k, n+1]$, thus the statement holds. In the case (ii) where $n_k = n$ and $m > 1$, we have

$$P = \lambda x x_2 \ldots x_l.\ x\ x_2 \ldots x_n\ x_{n+1}\ (e_1\ e_2)\ e_3 \ldots e_m.$$

whose corresponding binary tree t is $\langle\langle\langle \ldots \langle\langle \star, \underbrace{\star\rangle, \star\rangle, \ldots, \star\rangle}_{n \text{ leaves}}, \langle t_1, t_2\rangle, t_3\rangle, \ldots, t_m\rangle$. Hence, $\mathcal{L}(t) = \mathcal{L}(t') +\!\!+ [n+1]$ holds again from equation (2.3). In the case (iii) where $n_k = n + 1$, we have

$$P = \lambda x x_2 \ldots x_l.\ x\ x_2 \ldots x_n\ x_{n+1}\ (x_{n+2}\ e_1)\ e_2 \ldots e_m, \text{ or}$$

whose corresponding binary tree t is $\langle\langle\langle \ldots \langle\langle \star, \underbrace{\star\rangle, \star\rangle, \ldots, \star\rangle}_{n \text{ leaves}}, \langle \star, t_1\rangle, t_2\rangle, \ldots, t_m\rangle$. Hence, $\mathcal{L}(t) = \mathcal{L}(t') +\!\!+ [n+1]$ holds from equation (2.3). In the case (iv) where $n_k \geq n + 2$, we have

$$P = \lambda x x_2 \ldots x_l.\ x\ x_2 \ldots x_n\ x_{n+1}\ (x_{n+2}\ x_{n+3})\ \ldots\ e_1 \ldots e_m,$$

whose corresponding binary tree t is $\langle\langle\langle \ldots \langle\langle \star, \underbrace{\star\rangle, \star\rangle, \ldots, \star\rangle}_{n \text{ leaves}}, \langle \star, \star\rangle, \ldots, t_1\rangle, \ldots, t_m\rangle$. Hence, $\mathcal{L}(t) = \mathcal{L}(t') +\!\!+ [n+1]$ holds from equation (2.3). □

Example 2.9. Consider the λ-terms e_1, e_2, e_3 given in Figure 2. The λ-terms e_1, e_2, and e_3 given in Figure 2 are $\beta\eta$-equivalent to B^1B, $(B^3B) \circ (B^1B) \circ (B^1B)$, and $(B^5B) \circ (B^2B) \circ (B^2B) \circ (B^2B) \circ (B^0B)$, respectively, since $\mathcal{L}(t_1) = [1]$, $\mathcal{L}(t_2) = [3, 1, 1]$, $\mathcal{L}(t_3) = [5, 2, 2, 2, 0]$. (Recall t_j $(j = 1, 2, 3)$ is the unlabeled binary tree corresponding to e_j)

We conclude the uniqueness of decreasing polynomials for B-terms shown in the following theorem.

Theorem 2.10. *Every B-term e has a unique decreasing polynomial.*

Proof. For any given B-term e, we can find a decreasing polynomial for e from Lemma 2.3 and Lemma 2.4. Since no other decreasing polynomial can be equivalent to e from Lemma 2.8, the present statement holds. □

This theorem implies that the decreasing polynomial of B-terms can be used as their *canonical representation*, which is effectively derived as shown in Lemma 2.3 and Lemma 2.4.

As a corollary of the theorem, we can show the "only if" statement of Theorem 2.1, which corresponds to the completeness of the equation system.

Proof of Theorem 2.1. Let e_1 and e_2 be equivalent B-terms, that is, their λ-terms are $\beta\eta$-equivalent. From Theorem 2.10, their decreasing polynomials are the same. Since the decreasing polynomial is derived from e_1 and e_2 by equations (B1), (B2), and (B3) according to the proofs of Lemma 2.3 and Lemma 2.4, equivalence between e_1 and e_2 is also derived from these equations. □

Comparison with Curry's compositive normal form. Curry [Cur30b] has introduced a similar normal form for terms built from regular combinators[1], including B-terms. Curry's normal form is called *compositive* [Pip89] since it is given as a composition of four special terms, a K-term, W-term, C-term, and B-term. A B-term in Curry's normal form is expressed by

$$(B^{n_1}B^{m_1}) \circ (B^{n_2}B^{m_2}) \circ \cdots \circ (B^{n_k}B^{m_k})$$

where $k > 0$, $n_1 > n_2 > \cdots > n_k \geq 0$ and $m_i > 0$ for any $i = 1, \ldots, k$. Since we have

$$B^nB^m = B^n(\underbrace{B \circ \cdots \circ B}_{m}) = \underbrace{(B^nB) \circ \cdots \circ (B^nB)}_{m}$$

because of equation (B2'), the form is equivalent to

$$\underbrace{(B^{n_1}B) \circ \cdots \circ (B^{n_1}B)}_{m_1} \circ \underbrace{(B^{n_2}B) \circ \cdots \circ (B^{n_2}B)}_{m_2} \circ \cdots \circ \underbrace{(B^{n_k}B) \circ \cdots \circ (B^{n_k}B)}_{m_k}$$

which gives a decreasing polynomial. Curry informally proved the uniqueness of the normal form by an observation that $B^nB^m = \lambda x_0. \ldots x_{n+m+1}.x_0 \ldots x_n\ (x_{n+1} \ldots x_{n+m+1})$, while we have shown the exact correspondence between a B-terms as a lambda term and its normal form in decreasing polynomial representation.

2.4. **Relationship with Thompson's Group.** In this subsection, we explore a relationship between polynomials and *Thompson's group* F [MT73]. Thompson's group F is defined to be the group generated by formal elements x_n ($n = 0, 1, \ldots$) with relations $x_mx_n = x_nx_{m+1}$ for any $m > n$. Consider the map

$$f : \mathbf{CL}(B) \ni (B^{n_1}B) \circ \cdots \circ (B^{n_k}B) \mapsto x_{n_1}^{-1} \ldots x_{n_k}^{-1} \in F.$$

The map f is well-defined since for any $m > n$,

$$f((B^nB)\circ(B^mB)) = x_n^{-1}x_m^{-1} = (x_mx_n)^{-1} = (x_nx_{m+1})^{-1} = x_{m+1}^{-1}x_n^{-1} = f((B^{m+1}B)\circ(B^nB)).$$

We can think of $(\mathbf{CL}(B), \circ)$ as a semigroup since $(X \circ Y) \circ Z = X \circ (Y \circ Z)$ for any $X, Y, Z \in \mathbf{CL}(B)$, and $f : \mathbf{CL}(B) \to F$ is a semigroup homomorphism under this semigroup structure of $\mathbf{CL}(B)$. By definition, f is a semigroup isomorphism between $\mathbf{CL}(B)$ and the subsemigroup N of F generated by x_n^{-1} ($n = 0, 1, \ldots$).

It is known [Bel04] that every element of N corresponds to an infinite sequence of binary trees $(t_0, t_1, \ldots)$ (called a binary forest) where there exists k_0 such that $t_k = \star$ for any $k \geq k_0$.

Definition 2.11. The binary forest representation of an element of N is defined inductively as follows.

(1) The binary forest representation of x_n^{-1} is $(\underbrace{\star, \ldots, \star}_{n}, \langle\star, \star\rangle, \star, \ldots)$.

(2) If $y \in N$ corresponds to the binary forest $(t_0, t_1, \ldots)$, yx_n^{-1} corresponds to the binary forest

$$(t_0, t_1, \ldots, t_{n-1}, \langle t_n, t_{n+1}\rangle, t_{n+2}, \ldots).$$

[1] A regular combinator is a combinator in which no lambda abstraction occurs inside function application.

We can see the binary forests corresponding to $x_n^{-1}x_m^{-1}$ and $x_{m+1}^{-1}x_n^{-1}$ are equal to each other for any n, m.

(In fact, [Bel04] gave forest representations for the elements in the submonoid of F generated by x_n $(n = 0, 1, \dots)$, not x_n^{-1}). We show the binary forest representation of $x_{n_1}^{-1} \dots x_{n_k}^{-1}$ can be obtained from the binary tree corresponding to the λ-term of $(B^{n_1}B) \circ \dots \circ (B^{n_k}B)$.

Theorem 2.12. *Let $\langle \dots \langle \langle \star, t_1 \rangle, t_2 \rangle \dots, t_k \rangle$ be the binary tree corresponding to the λ-term of the polynomial $(B^{n_1}B) \circ \dots \circ (B^{n_k}B)$. Then, the binary forest representation of $f((B^{n_1}B) \circ \dots \circ (B^{n_k}B)) = x_{n_1}^{-1} \dots x_{n_k}^{-1}$ is given by*

$$(t_1, t_2, \dots, t_k, \star, \star, \dots).$$

Proof. We prove the theorem by induction on k. For binary trees $t_1, t_2, \dots, t_m$, we write $\varphi(t_1, t_2, \dots, t_m)$ for the binary tree $\langle \dots \langle \langle \star, t_1 \rangle, t_2 \rangle, \dots, t_m \rangle$. Since the binary tree corresponding to the λ-term of B^nB is given by $\varphi(\underbrace{\star, \dots, \star}_{n}, \langle \star, \star \rangle)$, the statement holds for the binary forest representations of $x_n = f(B^nB)$. Suppose $n_1 \geq \dots \geq n_k \geq n_{k+1}$. Then, the binary forest representation of $x_{n_1}^{-1} \dots x_{n_k}^{-1} x_{n_{k+1}}^{-1}$ is in the form of $(\underbrace{\star, \dots, \star}_{n_{k+1}}, \langle t_1, t_2 \rangle, t_3, \dots, t_m, \star, \dots)$. The binary tree $t = \varphi(\underbrace{\star, \dots, \star}_{n_{k+1}}, \langle t_1, t_2 \rangle, t_3, \dots, t_m)$ satisfies $\mathcal{L}(t) = [n_1, \dots, n_k, n_{k+1}]$ if the binary tree $t' = \varphi(\underbrace{\star, \dots, \star}_{n_{k+1}}, t_1, t_2, t_3, \dots, t_m)$ satisfies $\mathcal{L}(t') = [n_1, \dots, n_k]$. By Lemma 2.8, t is the binary tree corresponding to the λ-term of $(B^{n_1}B) \circ \dots \circ (B^{n_{k+1}}B)$, and this implies the desired result. $\square$

3. Results on the ρ-property of B-terms

In this section we show several approaches to if- and only-if-parts of Conjecture 1.1 for their special cases. For B-terms having the ρ-property, we introduce an efficient implementation to compute the entry point and the size of the cycle. For B-terms not having the ρ-property, we give two methods for proving that they do not have it.

3.1. B-terms having the ρ-property.

As shown in Section 1, we can check that B-terms equivalent to B^nB with $n \leq 6$ have the ρ-property by computing $(B^nB)_{(i)}$ for each i. However, it is not easy to check it by computer without an efficient implementation because we should compute all $(B^6B)_{(i)}$ with $i \leq 2980054085040$ $(= 2641033883877 + 339020201163)$ to know $\rho(B^6B) = (2641033883877, 339020201163)$. A naive implementation which computes terms of $(B^6B)_{(i)}$ for all i and stores all of them has no hope to detect the ρ-property.

We introduce an efficient procedure to find the ρ-property of B-terms which can successfully compute $\rho(B^6B)$. The procedure is based on two orthogonal ideas, Floyd's cycle-finding algorithm [Knu97] and an efficient right application algorithm over decreasing polynomials presented in Section 2.3.

The first idea, Floyd's cycle-finding algorithm (also called the tortoise and the hare algorithm), enables us to detect the cycle with constant memory usage, that is, the history of all terms $X_{(i)}$ does not need to be stored to check the ρ-property of the X combinator. The key to this algorithm is the fact that there are two distinct integers i and j with $X_{(i)} = X_{(j)}$ if and only if there is an integer m with $X_{(m)} = X_{(2m)}$, where the latter requires to compare $X_{(i)}$ and $X_{(2i)}$ from smaller i and store only these two terms for the next comparison between $X_{(i+1)} = X_{(i)}X$ and $X_{(2i+2)} = X_{(2i)}XX$ when $X_{(i)} \neq X_{(2i)}$. The following procedure computes the entry point and the size of the cycle if X has the ρ-property.

(1) Find the smallest m such that $X_{(m)} = X_{(2m)}$.
(2) Find the smallest k such that $X_{(k)} = X_{(m+k)}$.
(3) Find the smallest $0 < c \leq k$ such that $X_{(m)} = X_{(m+c)}$. If not found, put $c = m$.

After this procedure, we find $\rho(X) = (k, c)$. The third step can be run in parallel during the second one. See [Knu97, exercise 3.1.6] for the detail. Although we have tried the other cycle detection algorithm developed by Brent [Bre80] and Gosper [BGS72, item 132], they show a similar performance.

Efficient cycle-finding algorithms do not suffice to compute $\rho(B^6B)$. Only with the idea above running on a laptop (2.7 GHz Intel Core i7 / 16GB of memory), it takes about 2 hours even for $\rho(B^5B)$ and fails to compute $\rho(B^6B)$ with an out-of-memory error.

The second idea enables us to compute $X_{(i+1)}$ efficiently from $X_{(i)}$ for B-terms X. The key to this algorithm is to use the canonical representation of $X_{(i)}$, that is a decreasing polynomial, and directly compute the canonical representation of $X_{(i+1)}$ from that of $X_{(i)}$. Additionally, the canonical representation enables us to quickly decide equivalence which is required many times to find the cycle. It takes time just proportional to their lengths. If the λ-terms are used for finding the cycle, both application and deciding equivalence require much more complicated computation. Our implementation based on these two ideas computes $\rho(B^5B)$ and $\rho(B^6B)$ in 2 minutes and 6 days, respectively.

For two given decreasing polynomials P_1 and P_2, we show how a decreasing polynomial P equivalent to $(P_1\ P_2)$ can be obtained. The method is based on the following lemma about an application of one B-term to another B-term.

Lemma 3.1. *For B-terms e_1 and e_2, there exists $k \geq 0$ such that $e_1 \circ (B\ e_2) = B\ (e_1\ e_2) \circ B^k$.*

Proof. Let P_1 be a decreasing polynomial equivalent to e_1. We prove the statement by case analysis on the maximum degree in P_1. When the maximum degree is 0, we can take $k' \geq 1$ such that $P_1 \equiv \underbrace{B \circ \cdots \circ B}_{k'} = B^{k'}$. Then,

$$e_1 \circ (B\, e_2) = \underbrace{B \circ \cdots \circ B}_{k'} \circ (B\ e_2) = (B^{k'+1}\, e_2) \circ \underbrace{B \circ \cdots \circ B}_{k'} = B\,(e_1\, e_2) \circ B^{k'}$$

where equation (B3') is used k' times in the second equation. Therefore the statement holds by taking $k = k'$. When the maximum degree is greater than 0, we can take a decreasing polynomial P' for a B-term and $k' \geq 0$ such that $P_1 = (B\, P') \circ \underbrace{B \circ \cdots \circ B}_{k'} = (B\, P') \circ B^{k'}$

due to equation (B2'). Then,

$$\begin{aligned}
e_1 \circ (B\, e_2) &= (B\, P') \circ \underbrace{B \circ \cdots \circ B}_{k'} \circ (B\, e_2) \\
&= (B\, P') \circ (B^{k'+1}\, e_2) \circ \underbrace{B \circ \cdots \circ B}_{k'} \\
&= B\ (P' \circ (B^{k'}\, e_2)) \circ B^{k'} \\
&= B\ (B\, P'\, (B^{k'}\, e_2)) \circ B^{k'} \\
&= B\ (P_1\, e_2) \circ B^{k'} \\
&= B\ (e_1\, e_2) \circ B^{k'}.
\end{aligned}$$

Therefore, the statement holds by taking $k = k'$. □

This lemma indicates that, from two decreasing polynomials P_1 and P_2, a decreasing polynomial P equivalent to $(P_1\, P_2)$ can be obtained in the following steps where L_1 and L_2 are lists of non-negative numbers as shown in Section 2.3 corresponding to P_1 and P_2.

Algorithm 3.2 (Application of B-terms P_1 and P_2 in canonical representation).

(1) Build P'_2 by raising each degree of P_2 by 1, i.e., when $P_2 \equiv (B^{n_1}\, B) \circ \cdots \circ (B^{n_l}\, B)$, $P'_2 \equiv (B^{n_1+1}\, B) \circ \cdots \circ (B^{n_l+1}\, B)$. In terms of the list representation, a list L'_2 is built from L_2 by incrementing each element.
(2) Find a decreasing polynomial P_{12} corresponding to $P_1 \circ P'_2$ by equation (2.2). In terms of the list representation, a list L_{12} is constructed by appending L_1 and L'_2 and repeatedly applying (2.2).
(3) Obtain P by lowering each degree of P_{12} after eliminating the trailing 0-degree units, i.e., when $P_{12} \equiv (B^{n_1}\, B) \circ \cdots \circ (B^{n_l}\, B) \circ (B^0\, B) \circ \cdots \circ (B^0\, B)$ with $n_1 \geq \cdots \geq n_l > 0$, $P \equiv (B^{n_1-1}\, B) \circ \cdots \circ (B^{n_l-1}\, B)$. In terms of the list representation, a list L is obtained from L_{12} by decrementing each element after removing trailing 0's.

In the first step, a decreasing polynomial P'_2 equivalent to $B\, P_2$ is obtained. The second step yields a decreasing polynomial P_{12} for $P_1 \circ P'_2 = P_1 \circ (B\, P_2)$. Since P_1 and P_2 are decreasing, it is easy to find P_{12} by repetitive application of equation (2.2) for each unit of P'_2, à la insertion operation in insertion sort. In the final step, a polynomial P that satisfies $(B\, P) \circ B^k = P_{12}$ with some k is obtained. From Lemma 3.1 and the degree of decreasing polynomials, P is equivalent to $(P_1\, P_2)$.

Example 3.3. Let P_1 and P_2 be decreasing polynomials represented by lists $L_1 = [4, 1, 0]$ and $L_2 = [2, 0]$. Then a decreasing polynomial P equivalent to $(P_1\, P_2)$ is obtained as a list L in three steps:

(1) A list $L'_2 = [3, 1]$ is obtained from L_2.
(2) A decreasing list L_{12} is obtained by

$$L_{12} = [4, 1, \underline{0, 3}, 1] = [4, \underline{1, 4}, 0, 1] = [\underline{4, 5}, 1, 0, 1] = [6, 4, 1, \underline{0, 1}] = [6, 4, \underline{1, 2}, 0] = [6, 4, 3, 1, 0]$$

where equation (2.2) is applied in each underlined pair.
(3) A list $L = [5, 3, 2, 0]$ is obtained from L_{12} as the result of the application by decrementing each element after removing trailing 0's.

The implementation based on the right application over decreasing polynomials is available at `https://github.com/ksk/Rho` as a program named `bpoly`. In the current implementation, every decreasing polynomial is represented by a list (simulated by an array with offset and live length) whose k-th element stores the number of occurrences of (B^kB). For example, $(B^3B)\circ(B^2B)\circ(B^2B)\circ(B^0B)\circ(B^0B)$ is represented by a list `[2,0,2,1]` where the 0-th element is the leftmost 2. Since $(B^kB)^m$ is equivalent to B^kB^m, this representation can be seen as a variant of Curry's normal form mentioned in Section 2.3 by inserting the identity function B^kB^0 for each skipped degree k (*e.g.*, $(B^3B^1)\circ(B^2B^2)\circ(B^1B^0)\circ(B^0B^2)$ for the above). Using the Curry's normal form, we can adopt a slightly-improved algorithm by equation $B^nB^m\circ B^{n'}B^{m'} = B^{n'+m}B^{m'}\circ B^nB^m$ if $n<n'$ at the step (2) in Algorithm 3.2. Regarding cycle detection of the implemenation, Floyd's, Brent's and Gosper's algorithms are used. Note that the program does not terminate for the combinator which does not have the ρ-property. It will not help to decide if a combinator has the ρ-property. One might observe how the terms grow by repetitive right applications through running the program, though.

3.2. **B-terms not having the ρ-property.** A computer can check that a B-term has the ρ-property just by calculation but cannot show that a B-term does not have the ρ-property. In this subsection, we present two methods to prove that specific B-terms do not have the ρ-property. One employs decreasing polynomial representation as previously discussed and the other makes use of tree grammars for binary tree representation.

3.2.1. *Using polynomial representation.* We show that B^2 does not have ρ-property as an experiment. Note that B^2 has the decreasing polynomial representation $(B^0B)\circ(B^0B)$ which has the smallest length, 2, among the B-terms that are expected not to have the ρ-property.

To disprove the ρ-property of B^2, we show the following lemmas about the regularity of decreasing polynomial representation of $B^2_{(i)}$ for certain i. In these statements, we use

$$t_m = \frac{m^2+m}{2} \qquad \text{and} \qquad \bigodot_{i=k}^{n} f_i = f_k\circ f_{k+1}\circ f_{k+2}\circ\cdots\circ f_{n-1}\circ f_n.$$

In particular, $\bigodot_{i=k}^{n} f_i$ is an identity function if $k>n$.

Lemma 3.4. *For any k and m with $0\le k\le m$ and $l>0$,*

$$\bigodot_{i=k}^{m}(B^{m-i}\,B)^2\circ(B^l\,B)^2 = (B^{2m-2k+l+2}\,B)^2\circ\bigodot_{i=k}^{m}(B^{m-i}\,B)^2 \tag{3.1}$$

holds.

Proof. This statement can be obtained by applying equation (2.2) for $4(m-k+1)$ times. □

Lemma 3.5. *For any $m\ge 1$ and $0\le j\le m$,*

$$B^2_{(t_m+j)} = \bigodot_{i=1}^{j}(B^{2m-i-j+2}\,B)^2\circ\bigodot_{i=j+1}^{m}(B^{m-i}\,B)^2 \tag{3.2}$$

holds.

Proof. We prove the statement by induction on m. In the case of $m = 1$, $t_m = 1$. When $j = 0$, equation (3.2) is clear. When $j = 1$, equation (3.2) is shown by

$$\begin{aligned} B^2_{(2)} &= ((B^0\ B) \circ (B^0\ B))\ ((B^0\ B) \circ (B^0\ B)) \\ &= (B^2\ B) \circ (B^2\ B) = (B^2\ B)^2 \end{aligned}$$

by the application procedure over decreasing polynomial representation.

For the step case, we show that if equation (3.2) holds for $m = k \geq 1$ and $0 \leq j \leq k$, then it also holds for $m = k+1$ and $0 \leq j \leq k+1$. It is proved by induction on j where k is fixed. When $j = 0$, from the outer induction hypothesis, we obtain

$$\begin{aligned} B^2_{(t_{k+1})} &= B^2_{(t_k+k+1)} \\ &= B^2_{(t_k+k)}\, B^2 \\ &= \left(\bigodot_{i=1}^{k} (B^{2k-i-k+2}\, B)^2 \right) \ \big((B^0\, B) \circ (B^0\, B)\big) \\ &= \bigodot_{i=1}^{k} (B^{k-i+1}\, B)^2 \circ \big(B^0\, B\big) \circ \big(B^0\, B\big) \\ &= \bigodot_{i=1}^{k+1} (B^{(k+1)-i}\, B)^2 \end{aligned}$$

by applying the application procedure over decreasing polynomial representations, hence the statement holds for $j = 0$. When $0 < j \leq k+1$, from the inner induction hypothesis and Lemma 3.4, we similarly obtain

$$\begin{aligned} B^2_{(t_{k+1}+j)} &= B^2_{(t_{k+1}+j-1)}\, B^2 \\ &= \left(\bigodot_{i=1}^{j-1} (B^{2k-i-j+5}\, B)^2 \circ \bigodot_{i=j}^{k+1} (B^{k-i+1}\, B)^2 \right) \big((B^0\, B) \circ (B^0\, B)\big) \\ &= \bigodot_{i=1}^{j-1} (B^{2k-i-j+4}\, B)^2 \circ \bigodot_{i=j}^{k} (B^{k-i}\, B)^2 \circ \big(B^2\, B\big) \circ \big(B^2\, B\big) \\ &= \bigodot_{i=1}^{j-1} (B^{2k-i-j+4}\, B)^2 \circ \Big(B^{2k-2j+4}\, B\Big)^2 \circ \bigodot_{i=j}^{k} (B^{k-i}\, B)^2 \\ &= \bigodot_{i=1}^{j} (B^{2(k+1)-i-j+2}\, B)^2 \circ \bigodot_{i=j+1}^{k+1} (B^{(k+1)-i}\, B)^2. \end{aligned}$$

Therefore, the statement holds for $m = k+1$. □

These lemmas immediately lead to the anti-ρ-property of B^2.

Theorem 3.6. *The B-term B^2 does not have the ρ-property.*

Proof. We prove the statement by contradiction. If B^2 has the ρ-property, then the set of the normal forms of $S = \{B^2_{(i)} \mid i > 0\}$ is finite. Hence we can take m as the maximum length of decreasing polynomial representation among all B-terms in S. However, decreasing polynomial representation of $B^2_{(t_{m+1})}$ has length $m+1$ according to Lemma 3.5. This contradicts the assumption of m. $\square$

3.2.2. *Using tree grammars.* Another way for disproving the ρ-property of B-terms is to consider the $\beta\eta$-normal form of their λ-terms. As shown in Section 2, the $\beta\eta$-normal form of a B-term can be regarded as a binary tree. We can disprove the ρ-property of B-terms by observing what happens on the binary trees during the repetitive right application. More specifically, we first find a set which is closed under the application of a given term, and then show the length of the spine of trees is unbounded on the repetitive right application. This leads to the anti-ρ-property of the term as shown in Theorem 3.8.

First, we introduce some notations. In this section, we write $\langle\star,\star,\star,\dots,\star\rangle$ for the binary tree $\langle\dots\langle\langle\star,\star\rangle,\star\rangle,\dots,\star\rangle$ and identify B-terms with their corresponding binary trees. For a binary tree $t=\langle\star,t_1,\dots,t_k\rangle$, we define $l(t)=$ (the number of leaves in t), $a(t)=k$, and $N_i(t)=t_i$ for $i=1,\dots,k$. If X' is a B-term, $l(X')$, $a(X')$, and $N_i(X')$ are defined to be $l(t)$, $a(t)$, and $N_i(t)$ for the binary tree t corresponding to X'. Suppose the $\beta\eta$-normal form of X' is $\lambda x'_1\dots x'_{n'}.\,x'_1\ e_1\ \dots\ e_k$ and let X be another B-term whose $\beta\eta$-normal form is $\lambda x_1\dots x_n.\,e$. We can see $X'\ X=(\lambda x'_1\dots x'_{n'}.\,x'_1\ e_1\ \cdots\ e_k)\ X=\lambda x'_2\dots x'_{n'}.\,X\ e_1\ \cdots\ e_k$ and from Lemma 2.6, its $\beta\eta$-normal form is

$$\begin{cases}\lambda x'_2\dots x'_{n'}x_{k+1}\dots x_n.\,e[e_1/x_1,\dots,e_k/x_k] & (k\le n)\\ \lambda x'_2\dots x'_{n'}.\,e[e_1/x_1,\dots,e_n/x_n]\ e_{n+1}\ \cdots\ e_k & (\text{otherwise}).\end{cases}$$

Here $e[e_1/x_1,\dots,e_k/x_k]$ is the term which is obtained by substituting $e_1,\dots,e_k$ to the variables $x_1,\dots,x_k$ in e.

By simple computation with this fact, we get the following lemma:

Lemma 3.7. *Let X and X' be B-terms. Then*

$$l(X'\ X)=l(X')-1+\max\{l(X)-a(X'),\ 0\} \tag{3.3}$$

$$a(X'\ X)=a(X)+a(N_1(X'))+\max\{a(X')-l(X),\ 0\} \tag{3.4}$$

$$N_1(X'\ X)=\begin{cases}N_1(X)[N_2(X')/x_2,\dots,N_m(X')/x_m] & (\textit{if } N_1(X') \textit{ is a leaf})\\ N_1(N_1(X')) & (\textit{otherwise})\end{cases} \tag{3.5}$$

where $m=\min\{l(X),\ a(X')\}$.

From this lemma, we obtain a key theorem to prove the anti-ρ-property.

Theorem 3.8. *Let X be a B-term and T be a set of B-terms. If $\{X_{(i)}\mid i\ge 1\}\subset T$ and $l(X)-a(X')\ge 1$ for any $X'\in T$, then X does not have the ρ-property.*

Proof. It suffices to show the following: Under the hypotheses of the theorem, for any $i\ge 1$, there exists $j>i$ that satisfies $l(X_{(j)})>l(X_{(i)})$. Suppose, for contradiction, that there exists $i\ge 1$ that satisfies $l(X_{(i)})=l(X_{(j)})$ for any $j>i$. We get $a(X_{(j)})=l(X)-1$ by (3.3) and then $a(N_1(X_{(j-1)}))=l(X)-a(X)-1$ by (3.4). Here, $l(X)-a(X)\ge 2$ since if the $\beta\eta$-normal form of X is $\lambda x'_1\dots x'_{n'}.\,x'_1\ e_1\ \dots\ e_k$, each e_i $(i=1,\dots,k-1)$ has at least one variable and e_k has at least two variables because otherwise the λ-term is not η-normal. Therefore $a(N_1(X_{(j-1)}))\ge 1$, so $N_1(X_{(j-1)})$ is not a leaf for any $j>i$. From (3.5), we obtain $N_1(X_{(j-1)}) = N_1(N_1(X_{(j-2)})) = \cdots = \underbrace{N_1(\cdots N_1(}_{j-i} X_{(i)}) \cdots)$ for any $j > i$. However, this implies that $X_{(i)}$ has infinitely many variables and it yields contradiction. $\square$

Using this theorem, we prove that the B-term $(B^kB)^{(k+2)n}$ does not have the ρ-property. The $\beta\eta$-normal form of $(B^kB)^{(k+2)n}$ is given by

$$\lambda x_1 \ldots x_{k+(k+2)n+2}.\, x_1\ x_2\ \cdots\ x_{k+1}\ (x_{k+2}\ x_{k+3}\ \cdots\ x_{k+(k+2)n+2}).$$

This is deduced from Lemma 2.8 since the binary tree corresponding to the above λ-term is $t = \langle \underbrace{\star, \ldots, \star}_{k+1}, \langle \underbrace{\star, \ldots, \star}_{(k+2)n+1} \rangle\rangle$ and $\mathcal{L}(t) = [\underbrace{k, \ldots, k}_{(k+2)n}]$. In particular, we get $l((B^kB)^{(k+2)n}) = k + (k+2)n + 2$.

To apply Theorem 3.8, we introduce a set $T_{k,n}$ which satisfies the hypotheses of Theorem 3.8. First we inductively define a set of terms $T'_{k,n}$ as follows:

(1) $\star \in T'_{k,n}$
(2) $\langle \star,\ s_1,\ \ldots,\ s_{(k+2)n} \rangle \in T'_{k,n}$ if $s_i = \star$ for a multiple i of $k+2$ and $s_i \in T'_{k,n}$ for the others.

Then we define $T_{k,n}$ by $T_{k,n} = \left\{ \langle t_0,\ t_1,\ \ldots,\ t_{k+1} \rangle \,\middle|\, t_0,\ t_1, \ldots, t_{k+1} \in T'_{k,n} \right\}$. Since the binary tree of $(B^kB)^{(k+2)n}$ is $\langle \underbrace{\star, \ldots, \star}_{k+1}, \langle \star, \underbrace{\star, \ldots, \star}_{(k+2)n} \rangle\rangle$, we can see $(B^kB)^{(k+2)n} \in T_{k,n}$. Now we shall prove that $T_{k,n}$ is closed under right application of $(B^kB)^{(k+2)n}$.

Lemma 3.9. *If $X \in T_{k,n}$ then $X\ (B^kB)^{(k+2)n} \in T_{k,n}$.*

Proof. From the definition of $T_{k,n}$, if $X \in T_{k,n}$ then X can be written in the form $\langle t_0,\ t_1,\ \ldots,\ t_{k+1} \rangle$ for some $t_0, \ldots, t_{k+1} \in T'_{k,n}$. In the case where $t_0 = \star$, we have $X\ (B^kB)^{(k+2)n} = \langle t_1,\ \ldots,\ t_{k+1},\ \langle \star,\ \underbrace{\star,\ \ldots,\ \star}_{(k+2)n} \rangle\rangle \in T_{k,n}$. In the case where t_0 has the form of 2 in the definition of $T'_{k,n}$, then we have $X = \langle \star,\ s_1,\ \ldots,\ s_{(k+2)n},\ t_1,\ \ldots,\ t_{k+1} \rangle$ with $s_i = \star$ for a multiple i of $k+2$ and $s_i \in T'_{k,n}$ for others, hence

$$X\ (B^kB)^{(k+2)n} = \langle s_1,\ \ldots,\ s_{k+1},\ \langle s_{k+2},\ \ldots,\ s_{(k+2)n},\ t_1,\ \ldots,\ t_{k+1},\ \star \rangle\rangle.$$

We can easily see $s_1,\ \ldots,\ s_{k+1}$, and $\langle s_{k+2},\ \ldots,\ s_{(k+2)n},\ t_1,\ \ldots,\ t_{k+1},\ \star \rangle$ are in $T'_{k,n}$. $\square$

From the definition of $T_{k,n}$, we can compute that $a(X)$ equals $k+1$ or $(k+2)n+k+1$ if $X \in T_{k,n}$. Particularly, we get the following:

Lemma 3.10. *For any $X \in T_{k,n}$, $a(X) \leq (k+2)n + k + 1 = l((B^kB)^{(k+2)n}) - 1$.*

By Theorem 3.8, we get the desired result:

Theorem 3.11. *For any $k \geq 0$ and $n > 0$, $(B^kB)^{(k+2)n}$ does not have the ρ-property.*

We give more examples of B-terms which satisfy the condition in Theorem 3.8 with some set T.

Example 3.12. Consider $X = (B^2B)^2 \circ (BB)^2 \circ B^2 = \langle \star,\ \langle \star,\ \langle \star,\ \langle \star,\ \star,\ \star \rangle,\ \star \rangle,\ \star \rangle\rangle$. We inductively define T' as follows:

(1) $\star \in T'$

(2) For any $t \in T'$, $\langle \star,\ t,\ \star \rangle \in T'$
(3) For any $t_1, t_2 \in T'$, $\langle \star,\ t_1,\ \star,\ \langle \star,\ t_2,\ \star \rangle,\ \star \rangle \in T'$
Then $T = \{\langle t_1,\ \langle \star,\ t_2,\ \star \rangle \rangle \mid t_1, t_2 \in T'\}$ satisfies the conditions in Theorem 3.8:

Claim 3.13. $\{X_{(k)} \mid k \geq 1\} \subset T$.

Proof. By definition, $X \in T$. Let $X' = \langle t_1,\ \langle \star,\ t_2,\ \star \rangle \rangle \in T$. Then, we have

$$X'X = \begin{cases} \langle \star,\ t_2,\ \star,\ \langle \star,\ \langle \star,\ \langle \star,\ \star,\ \star \rangle,\ \star \rangle,\ \star \rangle \rangle & \text{if } t_1 = \star \\ \langle t_1',\ \langle \star,\ \langle \star,\ t_2,\ \star,\ \langle \star,\ \star,\ \star \rangle,\ \star \rangle,\ \star \rangle \rangle & \text{if } t_1 = \langle \star,\ t_1',\ \star \rangle \\ \langle t_{11},\ \langle \star,\ \langle \star,\ t_{12},\ \star,\ \langle \star,\ \langle \star,\ t_2,\ \star \rangle,\ \star \rangle,\ \star \rangle,\ \star \rangle \rangle & \text{if } t_1 = \langle \star,\ t_{11},\ \star,\ \langle \star,\ t_{12},\ \star \rangle,\ \star \rangle, \end{cases}$$

and, in either case, $X'X \in T$. □

Claim 3.14. $l(X) - a(X') \geq 1$ *for any* $X' \in T$.

Proof. Since $a(X')$ is equal to either 1, 3, or 5, and $l(X) = 8$, $l(X) - a(X') \geq 3$. □

Thus, $(B^2B)^2 \circ (BB)^2 \circ B^2$ does not have the ρ-property.

Example 3.15. Consider $X = (BB)^3 \circ B^3 = \langle \star,\ \langle \star,\ \star,\ \star,\ \star \rangle,\ \star,\ \star \rangle$. We inductively define T' as follows:
(1) $\star \in T'$
(2) For any $t \in T'$, $\langle \star,\ t,\ \star,\ \star \rangle \in T'$
Then $T = \{\langle t_1, \langle \star, t_2, \star, \star \rangle \rangle \mid t_1, t_2 \in T'\}$ satisfies the conditions in Theorem 3.8:

Claim 3.16. $\{X_{(k)} \mid k \geq 1\} \subset T$.

Proof. By definition, $X \in T$. Let $X' = \langle t_1,\ \langle \star,\ t_2,\ \star,\ \star \rangle \rangle \in T$. Then, we have

$$X'X = \begin{cases} \langle \star,\ t_2,\ \star,\ \star,\ \langle \star,\ \langle \star,\ \star,\ \star,\ \star \rangle,\ \star,\ \star \rangle \rangle & \text{if } t_1 = \star \\ \langle t_1',\ \langle \star,\ \langle \star,\ \langle \star,\ t_2,\ \star,\ \star \rangle,\ \star,\ \star \rangle,\ \star,\ \star \rangle \rangle & \text{if } t_1 = \langle \star,\ t_1\ \star,\ \star \rangle \end{cases}$$

and, in either case, $X'X \in T$. □

Claim 3.17. $l(X) - a(X') \geq 1$ *for any* $X' \in T$.

Proof. $a(X')$ equals 1 or 4 and $l(X) = 8$, so $l(X) - a(X') \geq 4$. □

Thus, $(BB)^3 \circ B^3$ does not have the ρ-property.

Theorem 3.8 gives a possible technique to prove that $l(X_{(i)})$ diverges, or, the anti-ρ-property of X, for some B-term X. Since the hypotheses of Theorem 3.8 implies that $l(X_{(i)})$ is also monotonically non-decreasing, we can consider another problem on B-terms: "Give a necessary and sufficient condition for $l(X_{(i)})$ to be monotonically non-decreasing for a B-term X."

4. Possible approaches

The present paper introduces a canonical representation to make equivalence check of B-terms easier. The idea of the representation is based on that we can lift all $\circ$'s (2-argument B) to the outside of B (1-argument B) by equation (B2'). One may consider it the other way around. Using the equation, we can lift all B's (1-argument B) to the outside of $\circ$ (2-argument B). Then one of the arguments of $\circ$ becomes B. By equation (B3'), we can move all B's right. Thereby we find another canonical representation for B-terms given by

$$e \ ::= \ B \ \mid \ B\ e \ \mid \ e \circ B.$$

We can show the uniqueness of this representation by giving a bijective transformation f from it to the polynomial representation. We define f inductively by

$$\begin{aligned} f(B) &= B^0 B \\ f(B\ e) &= (B^{n_1+1}B) \circ \cdots \circ (B^{n_k+1}B) \quad \text{if } f(e) = (B^{n_1}B) \circ \cdots \circ (B^{n_k}B) \\ f(e \circ B) &= f(e) \circ (B^0 B). \end{aligned}$$

Note that $B^0 B = B$ and the second rule of f does not change the equivalence class of B-terms because $B(e_1 \circ \cdots \circ e_k) = (B\ e_1) \circ \cdots \circ (B\ e_k)$ (Equation (B2')). We can see the inverse of this function is given by

$$\begin{aligned} f^{-1}(B^0 B) &= B \\ f^{-1}((B^{n_1}B) \circ \cdots \circ (B^{n_k}B)) &= B\ (f^{-1}((B^{n_1-1}B) \circ \cdots \circ (B^{n_k-1}B))) \quad (n_k > 0) \\ f^{-1}((B^{n_1}B) \circ \cdots \circ (B^{n_k}B) \circ (B^0 B)) &= (f^{-1}((B^{n_1}B) \circ \cdots \circ (B^{n_k}B))) \circ B. \end{aligned}$$

Function application (written as @, explicitly) over this canonical representation can be recursively defined by

$$\begin{aligned} B \ @\ e &= B\ e \\ (e_1 \circ B) \ @\ e_2 &= e_1 \ @\ (B\ e_2) \\ (B\ e) \ @\ B &= e \circ B \\ (B\ e_1) \ @\ (e_2 \circ B) &= ((B\ e_1) \ @\ e_2) \circ B \\ (B\ B) \ @\ (B\ e) &= (B\ (B\ e)) \circ B \\ (B\ (e_1 \circ B)) \ @\ (B\ e_2) &= ((B\ e_1) \ @\ (B\ (B\ e_2))) \circ B \\ (B\ (B\ e_1)) \ @\ (B\ e_2) &= B\ ((B\ e_1) \ @\ e_2). \end{aligned}$$

Notice that the pattern matching is exhaustive. The correctness of the equations is proved by equations (B2') and (B3'). Termination of the recursive definition is shown by a simple lexicographical order of the first and the second operand of application. Note that this canonical form can be represented by a sequence of $(B\ \square)$ and $(\square \circ B)$ where $\square$ stands for a hole. Also, a sequence of them exactly corresponds to a single term in canonical form by hole application. *e.g.*, $[(B\ \square),\ (B\ \square),\ (\square \circ B)]$ represents $B\ (B\ (B \circ B))$ where a nullary constructor B is filled in the last element $(\square \circ B)$. This fact may be used to find the ρ- or anti-ρ-properties. By writing 0 and 1 for $(B\ \square)$ and $(\square \circ B)$, the above equation can be

Table 1. Summary of known results on the ρ-property of B-terms

having ρ-property	$B^n B$ with $0 \leq n \leq 6$
having anti-ρ-property	$(B^k B)^{(k+2)n}$ with $k \geq 0, n > 0$ $(B^2 B)^2 \circ (BB)^2 \circ B^2$ $(BB)^3 \circ B^3$

rewritten as follows:

$$\begin{aligned} \varepsilon @ y &= 0y \\ 1x @ y &= x @ 0y \\ 0x @ \varepsilon &= 1x \\ 0x @ 1y &= 1(0x @ y) \\ 0\varepsilon @ 0y &= 100y \\ 01x @ 0y &= 1(0x @ 00y) \\ 00x @ 0y &= 0(0x @ y) \end{aligned}$$

where ε is used for the end marker (filling B at the end). A monomial B-term corresponds to a binary sequence that does not contain 1. If $x@y$ is always greater than x in some measure when y contains 1, we can claim the "only-if" part of Conjecture 1.1.

Waldmann [Wal13] suggests that the ρ-property of $B^n B$ may be checked even without converting B-terms into canonical forms. He simply defines B-terms by

$$e ::= B^k \mid e\ e$$

and regards B^k as a constant which has a rewrite rule $B^k\ e_1\ e_2\ \ldots\ e_{k+2} \to e_1\ (e_2\ \ldots\ e_{k+2})$. He implemented a check program in Haskell to confirm the ρ-property. Even in the restriction on rewriting, he found that $(B^0 B)_{(9)} = (B^0 B)_{(13)}$, $(B^1 B)_{(36)} = (B^1 B)_{(56)}$, $(B^2 B)_{(274)} = (B^2 B)_{(310)}$ and $(B^3 B)_{(4267)} = (B^3 B)_{(10063)}$, in which it requires a few more right applications to find the ρ-property than the case of canonical representation. If the ρ-property of $B^n B$ for any $n \geq 0$ is shown under the restricted equivalence given by the rewrite rule, then we can conclude the "if" part of Conjecture 1.1.

5. Concluding remark

We have investigated the ρ-properties of B-terms in particular forms so far. Table 1 summarizes all results we obtained. While the B-terms equivalent to $B^n B$ with $n \leq 6$ have the ρ-property, the B-terms $(B^k B)^{(k+2)n}$ with $k \geq 0$ and $n > 0$, $(B^2 B)^2 \circ (BB)^2 \circ B^2$, and $(BB)^3 \circ B^3$ do not. We have also introduced a canonical representation of B-terms which is useful to prove or disprove of specific B-terms.

We introduce remaining problems related to these results. The ρ-property is defined for any combinatory terms (and closed λ-terms). We investigated it mainly for B-terms as a simple but interesting instance to give a partial solution of Conjecture 1.1 in the present paper. The conjecture implies that the ρ-property of B-terms is decidable. One could consider the decidability of the ρ-property for BCK and BCI-terms which is still open. Also, the decidability for the ρ-property of S-terms and L-terms can be considered. Waldmann's work on a rational representation of normalizable S-terms [Wal00] may be helpful to solve it. We expect that none of the S-terms have the ρ-property as S itself does not, though. Regarding L-terms, Statman's work [Sta89] may be helpful where equivalence of L-terms is shown decidable up to a congruence relation induced by $L\ e_1\ e_2 \to e_1\ (e_2\ e_2)$. It would be interesting to investigate the ρ-property of L-terms in this setting.

Acknowledgment

We are grateful to Johannes Waldmann and Sebastian Maneth for their fruitful comments on this work at an earlier stage. This work was partially supported by JSPS KAKENHI Grant Number JP17K00007.

References

[Bar84] Hendrik P. Barendregt. *The Lambda Calculus: Its Syntax and Semantics*. Studies in Logic and the Foundations of Mathematics. Elsevier Science, 1984.

[Bel04] James Belk. *Thompson's group F*. PhD thesis, Cornell University, 2004.

[BGS72] Michael Beeler, Ralph W. Gosper, and Richard C. Schroeppel. HAKMEM. Technical report, Massachusetts Institute of Technology, Cambridge, MA, USA, 1972.

[Bre80] Richard P. Brent. An improved Monte Carlo factorization algorithm. *BIT*, 20(2):176–184, 1980.

[Cur30a] Haskell B. Curry. Grundlagen der Kombinatorischen Logik (Teil I). *American Journal of Mathematics*, 52(3):509–536, 1930.

[Cur30b] Haskell B. Curry. Grundlagen der Kombinatorischen Logik (Teil II). *American Journal of Mathematics*, 52(4):789–834, 1930.

[IN18] Mirai Ikebuchi and Keisuke Nakano. On repetitive right application of b-terms. In *3rd International Conference on Formal Structures for Computation and Deduction, FSCD 2018, July 9-12, 2018, Oxford, UK*, pages 18:1–18:15, 2018.

[Knu97] Donald E. Knuth. *The Art of Computer Programming, Volume 2 (3rd Ed.): Seminumerical Algorithms*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1997.

[Lja68] Evgeny S. Ljapin. *Semigroups*. Translations of Mathematical Monographs. American Mathematical Society, 1968.

[MT73] Ralph McKenzie and Richard J. Thompson. An elementary construction of unsolvable word problems in group theory**this work was aided by national science foundation grants gp-7578 and gp-6232x3. In W.W. Boone, F.B. Cannonito, and R.C. Lyndon, editors, *Word Problems*, volume 71 of *Studies in Logic and the Foundations of Mathematics*, pages 457 – 478. Elsevier, 1973.

[Nak08] Keisuke Nakano. ρ-property of combinators. 29th TRS Meeting in Tokyo, 2008.

[Oka03] Chris Okasaki. Flattening combinators: surviving without parentheses. *Journal of Functional Programming*, 13(4):815–822, July 2003.

[Pip89] Adolfo Piperno. Abstraction problems in combinatory logic a compositive approach. *Theor. Comput. Sci.*, 66(1):27–43, 1989.

[Sch24] Moses Schönfinkel. Über die Bausteine der mathematischen Logik. *Mathematische Annalen*, 92(3–4):305–316, 1924.

[Smu12] Raymond M. Smullyan. *To Mock a Mockingbird*. Knopf Doubleday Publishing Group, 2012.

[Sta89] Rick Statman. The Word Problem for Smullyan's Lark Combinator is Decidable. *Journal of Symbolic Computation*, 7(2):103–112, February 1989.

[Sta11] Rick Statman. To Type A Mockingbird. Draft paper available from `http://tlca.di.unito.it/PAPER/TypeMock.pdf`, December 2011.

[Wal00] Johannes Waldmann. The Combinator S. *Information and Computation*, 159(1-2):2–21, May 2000.

[Wal13] Johannes Waldmann. Personal communication, March 2013.