Specification, Implementation, and Complexity of Replicated Data Types with Composite Operations

Hengfeng Wei
State Key Laboratory for Novel Software Technology, Nanjing University, China
hfwei@nju.edu.cn
https://orcid.org/0000-0002-0427-9710

Huang Yu
State Key Laboratory for Novel Software Technology, Nanjing University, China
yuhuang@nju.edu.cn
https://orcid.org/0000-0001-8921-036X

Jian Lu
State Key Laboratory for Novel Software Technology, Nanjing University, China
lj@nju.edu.cn

Abstract

2012 ACM Subject Classification

Keywords and phrases Replicated set

Digital Object Identifier 10.4230/LIPIcs...

1 Introduction

2 Preliminaries

2.1 Observed-Remove Set (OR-Set)

$$\mathcal{F}_{\texttt{orset}}(\texttt{rd}, E, \mathsf{op}, \mathsf{vis}, \mathsf{ar}) = \{a \mid \exists e \in E.\mathsf{op}(e) = \texttt{add}(a) \tag{1}$$

$$\land \big(\forall f \in E.\mathsf{op}(f) = \texttt{rm}(a) \implies \neg(e \xrightarrow{\mathsf{vis}} f)\big)\}. \tag{2}$$

3 Replicated Data Types with Composite Operations

3.1 Specification

We consider a composite operation of a replicated data type τ in the form of $C = A \oplus B$, where A, B, and C are different objects of type τ.

Q : Generalize to different data types?

Following [1], we specify the semantics of a composite operation $A \oplus B$ of a replicated data type τ by a function $\mathcal{F}_\tau$ that determines the return value of $\oplus$ based on prior operations performed on the two objects involved (i.e., A and B). However, $\mathcal{F}_\tau$ for a composite operation $\oplus$ takes as parameters two, not one as in [1], *operation contexts*, one on each object involved.

Note: Partial operation context [2]

▶ **Definition 1** (Product of Operation Contexts)**.** Consider two operation contexts for the same replicated data type τ:

$$\mathcal{C}_A = (E_A, \mathsf{op}_A, \mathsf{vis}_A, \mathsf{ar}_A) \tag{3}$$
$$\mathcal{C}_B = (E_B, \mathsf{op}_B, \mathsf{vis}_B, \mathsf{ar}_B) \tag{4}$$

The product $\mathcal{C} = \mathcal{C}_A \times \mathcal{C}_B$ of $\mathcal{C}_A$ and $\mathcal{C}_B$ is also an operation context defined as $\mathcal{C} = (E, \mathsf{op}, \mathsf{vis}, \mathsf{ar})$, where

- $E = E_A \times E_B$
- $\mathsf{op} = \mathsf{op}_A \sqcup \mathsf{op}_B$
- $\mathsf{vis} = \mathsf{vis}_A \times \mathsf{vis}_B$
- $\mathsf{ar} = \mathsf{ar}_A \times \mathsf{ar}_B$

▶ **Definition 2.**

$$\mathcal{F}_\tau(\oplus, \mathcal{C}_A, \mathcal{C}_B) = \mathcal{F}_\tau(\oplus, \mathcal{C}_A \times \mathcal{C}_B) \tag{5}$$

4 Replicated Counter with Composite Operations

We consider the replicated counter data type with composite operations including addition $(+)$ and subtraction $(-)$.

5 Replicated Set with Composite Operations

We consider the replicated set data type with composite operations including union $(\cup)$, intersection $(\cap)$, difference $(\setminus)$, and symmetric difference $(\otimes)$.

5.1 Specification

General method for set union, set intersection, and set difference: There are four $\cdots$

$$\mathcal{F}_{\texttt{orset}}(A \setminus B, \mathcal{C}_A, \mathcal{C}_B) \tag{6}$$
$$= \{a \mid \exists (e, e') \in E_A \times E_B.\Big(\big(\texttt{op}(e) = \texttt{add}(a) \land \texttt{op}(e') = \texttt{rm}(a)\big) \tag{7}$$
$$\land \forall (f, f') \in E_A \times E_B.\big((\texttt{op}(f) = \texttt{add}(a) \land \texttt{op}(f') = \texttt{add}(a)) \lor (\texttt{op}(f) = \texttt{rm}(a) \land \texttt{op}(f') = \texttt{add}(\tag{8}$$
$$\implies \neg\big((e, e') \xrightarrow{\mathsf{vis}} (f, f')\big)\Big)\}. \tag{9}$$

▶ **Example 3** (Specification for Composite Operations of Replicated Set)**.**

5.2 Protocol

5.2.1 Methodology

Op-based

We always assume that $A \neq \emptyset$, $A' \neq \emptyset$, and $(\mathbb{A} \cup \mathbb{R}) \cap (\mathbb{A}' \cup \mathbb{R}') = \emptyset$. Note that for any replica state $(a, \mathbb{A}, \mathbb{R})$, we have $\mathbb{R} \subseteq \mathbb{A}$.

Q : $\emptyset$ issues The goal is to find $\Box$ and $\Diamond$ such that

$$\cdots \iff \Box \subseteq \Diamond$$

$$
\begin{array}{lcl}
U & : & \text{universe of elements} \\
I & : & \text{set of identifiers} \\
\mathbb{A}, \mathbb{R} & \in & \text{SUBSET } \mathcal{I} \\
\Sigma & = & \mathcal{U} \times \mathbb{A} \times \mathbb{R} \\
\sigma_0 & = & \lambda r.\, \emptyset
\end{array}
$$

Figure 1 Implementation of replicated set with composite operations.

5.2.2 Basic Composite Operations for Replicated Set

5.2.2.1 Set Union ($\cup$)

The goal is to find $\Box$ and $\Diamond$ such that

$$(\mathbb{R} \subset \mathbb{A}) \vee (\mathbb{R}' \subset \mathbb{A}') \iff \Box \subset \Diamond.$$

▶ **Theorem 4** (Set Union). *Given that* $\mathbb{R} \subseteq \mathbb{A}$ *and* $\mathbb{R}' \subseteq \mathbb{A}'$*, we have*

$$(\mathbb{R} \subset \mathbb{A}) \vee (\mathbb{R}' \subset \mathbb{A}') \iff (\mathbb{R} \cup \mathbb{R}') \subset (\mathbb{A} \cup \mathbb{A}').$$

▶ **Theorem 5** (Set Union (Alternative)). *Given that* $\mathbb{R} \subseteq \mathbb{A}$ *and* $\mathbb{R}' \subseteq \mathbb{A}'$*, we have*

$$(\mathbb{R} \subset \mathbb{A}) \vee (\mathbb{R}' \subset \mathbb{A}') \iff (\mathbb{A} \times \mathbb{R}') \cap (\mathbb{R} \times \mathbb{A}') \subset (\mathbb{A} \times \mathbb{A}')$$

Q : How to choose? Another way:

$(\mathbb{R} \subset \mathbb{A}) \vee (\mathbb{R}' \subset \mathbb{A}$

Another way: IF THEN ELSE

5.2.2.2 Set Intersection ($\cap$)

The goal is to find $\Box$ and $\Diamond$ such that

$$(\mathbb{R} \subset \mathbb{A}) \wedge (\mathbb{R}' \subset \mathbb{A}') \iff \Box \subset \Diamond.$$

▶ **Theorem 6** (Set Intersection). *Given that* $\mathbb{R} \subseteq \mathbb{A}$ *and* $\mathbb{R}' \subseteq \mathbb{A}'$*, we have*

$$(\mathbb{R} \subset \mathbb{A}) \wedge (\mathbb{R}' \subset \mathbb{A}') \iff (\mathbb{A} \times \mathbb{R}') \cup (\mathbb{R} \times \mathbb{A}') \subset (\mathbb{A} \times \mathbb{A}').$$

5.2.2.3 Set Difference ($\setminus$)

The goal is to find $\Box$ and $\Diamond$ such that

$$(\mathbb{R} \subset \mathbb{A}) \wedge (\mathbb{R}' \not\subset \mathbb{A}') \iff \Box \subset \Diamond.$$

▶ **Theorem 7** (Set Difference). *Given that* $\mathbb{R} \subseteq \mathbb{A}$ *and* $\mathbb{R}' \subseteq \mathbb{A}'$*, we have*

$$(\mathbb{R} \subset \mathbb{A}) \wedge (\mathbb{R}' \not\subset \mathbb{A}') \iff (\mathbb{R} \times \mathbb{A}') \subset (\mathbb{A} \times \mathbb{R}').$$

5.2.3 Derived Composite Operations and Laws for Replicated Set

5.2.3.1 Set Symmetric Difference ($\otimes$)

The goal is to find $\Box$ and $\Diamond$ such that

$$\big((\mathbb{R} \subset \mathbb{A}) \wedge (\mathbb{R}' \not\subset \mathbb{A}')\big) \vee \big((\mathbb{R} \not\subset \mathbb{A}) \wedge (\mathbb{R}' \subset \mathbb{A}')\big) \iff \Box \subset \Diamond.$$

By the identity $A \otimes B = (A \cup B) \setminus (A \cap B)$ and Theorems 4, 6, and 7, we have

▶ **Theorem 8** (Set Symmetric Difference). *Given that* $\mathbb{R} \subseteq \mathbb{A}$ *and* $\mathbb{R}' \subseteq \mathbb{A}'$*, we have*

$$\big((\mathbb{R} \subset \mathbb{A}) \wedge (\mathbb{R}' \not\subset \mathbb{A}')\big) \vee \big((\mathbb{R} \not\subset \mathbb{A}) \wedge (\mathbb{R}' \subset \mathbb{A}')\big)$$
$$\iff (\mathbb{R} \cup \mathbb{R}') \times (\mathbb{A} \cup \mathbb{A}') \subset (\mathbb{A} \cup \mathbb{A}') \times \big((\mathbb{A} \times \mathbb{R}') \cup (\mathbb{R} \times \mathbb{A}')\big)$$

It fails to use Theorem 4 here.

By the identity $A \otimes B = (A \setminus B) \cup (B \setminus A)$ and Theorems 5 and 7, we have

▶ **Theorem 9** (Set Symmetric Difference (Alternative)). *Given that* $\mathbb{R} \subseteq \mathbb{A}$ *and* $\mathbb{R}' \subseteq \mathbb{A}'$*, we have*

$$\big((\mathbb{R} \subset \mathbb{A}) \wedge (\mathbb{R}' \not\subset \mathbb{A}')\big) \vee \big((\mathbb{R} \not\subset \mathbb{A}) \wedge (\mathbb{R}' \subset \mathbb{A}')\big)$$
$$\iff (\mathbb{A} \times \mathbb{R}') \times (\mathbb{R}' \times \mathbb{A}) \cap (\mathbb{R} \times \mathbb{A}') \times (\mathbb{A}' \times \mathbb{R}) \subset (\mathbb{A} \times \mathbb{R}') \times (\mathbb{A}' \times \mathbb{R})$$

5.2.3.2 Laws

▶ **Theorem 10** (Distributive Law).

$$A \cup (B \cap C) = (A \cup B) \wedge (A \cup C)$$

6 Related Work

[4] [3]

7 Conclusion and Future Work

References

1 Sebastian Burckhardt, Alexey Gotsman, Hongseok Yang, and Marek Zawirski. Replicated data types: Specification, verification, optimality. In *Proceedings of the 41st ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages*, POPL '14, pages 271–284, New York, NY, USA, 2014. ACM. URL: `http://doi.acm.org/10.1145/2535838.2535848`, `doi:10.1145/2535838.2535848`.

2 Alexey Gotsman and Hongseok Yang. Composite replicated data types. In *European Symposium on Programming Languages and Systems*, pages 585–609. Springer, 2015.

3 Radha Jagadeesan and James Riely. From sequential specifications to eventual consistency. In *Proceedings, Part II, of the 42Nd International Colloquium on Automata, Languages, and Programming - Volume 9135*, ICALP 2015, pages 247–259, Berlin, Heidelberg, 2015. Springer-Verlag. URL: `https://doi.org/10.1007/978-3-662-47666-6_20`, `doi:10.1007/978-3-662-47666-6_20`.

4 Christopher Meiklejohn and Peter Van Roy. Lasp: A language for distributed, coordination-free programming. In *Proceedings of the 17th International Symposium on Principles and Practice of Declarative Programming*, PPDP '15, pages 184–195, New York, NY, USA, 2015. ACM. URL: `http://doi.acm.org/10.1145/2790449.2790525`, `doi:10.1145/2790449.2790525`.