

Verifying Robustness of Programs Under Structural Perturbations

Jacob Bond and Clay Thomas

Purdue University

1 Introduction

The question of robustness is fundamental to the subject of programming by example (PBE). Robustness of a program is the property that the program behaves predictably on uncertain inputs [1]. In the PBE paradigm, there is, by definition, an uncertainty about the intent of the user, and therefore, it is desirable that a program synthesizer behave predictably with regards to this uncertainty.

Consider an attempt to specify the max function by providing to a program synthesizer the examples $(13, 15) \mapsto 15$, $(-23, 19) \mapsto 19$, and $(-75, -13) \mapsto -13$. In order to synthesize a simpler program, the result will be the program $P(a, b) := \text{return } b;$. The issue that arises here is that neither the program synthesizer, nor the synthesized program, are robust.

The program synthesizer is not robust, as transposing the inputs of each example would result in the program $P(a, b) := \text{return } a;$, while transposing just a single input would result in the correct program $P(a, b) := \text{return } a > b ? a : b;$.

Additionally, the program which is synthesized is not robust as $P(a, b) \neq P(b, a)$. That is, the program does not behave predictably under uncertainty in the order of the arguments. If the program which is synthesized is required to be robust with respect to uncertainty in the order of the input, neither $P(a, b) := \text{return } a;$ nor $P(a, b) := \text{return } b$ would be viable candidates, and the synthesizer would be forced to return $P(a, b) := \text{return } a > b ? a : b;$.

Moreover, a synthesizer which returns robust programs will itself be more robust. Let $\mathcal{I}_1 = (I_1, O_1)$ and $\mathcal{I}_2 = (I_2, O_2)$ be two input-output pairs for a program synthesizer which differ by a small perturbation. If the program \mathcal{P}_1 returned by the synthesizer on input \mathcal{I}_1 is robust, then $\mathcal{P}_1(I_2)$ will approximate O_2 because I_2 approximates I_1 . For this reason, \mathcal{P}_2 , the program returned on input \mathcal{I}_2 , should only differ from \mathcal{P}_1 by a small amount.

Thus, the issue of robustness in PBE can be addressed by verifying robustness, either of the synthesized programs or even of the synthesizer as a whole. However, verification of robustness requires the ability to reason about robustness. Standard program verifiers are unable to verify robustness properties as they are an example of a 2-safety property, a property which requires reasoning about two execution traces simultaneously. In particular, robustness is the property that given two inputs which are related by some form of uncertainty, the outputs should be related in a predictable way. As such, verifying robustness requires reasoning about the relationship between two execution traces.

2 Related Work

2.1 Early Work

Relational Logics Benton [2] established a Relational Hoare Logic in order to reason about relational properties, properties which consider two, usually distinct, programs. Barthe et al. [3, 4] extended Benton’s Relational Hoare Logic to probabilistic programs in order to reason about cryptographic protocols and differential privacy.

Self-Composition Barthe et al. [5] applied self-composition, sequential running of renamed copies of the original program, to study secure information flow. Terauchi & Aiken [6] built on this work by applying a type-based approach to complement self-composition.

2.2 Continuity and Lipschitz Robustness

Continuity Hamlet [9] considered the concept of program continuity, but declared that automating verification of continuity for programs with loops was infeasible. However, Chaudhuri et al. [10] was able to establish a program logic which allowed such automatic verification of continuity.

TODO: incorporate above and below. Below needs work to fit in this section

Robustness in the setting of numerical perturbations is realized by the property of continuity. In many situations, the result of a program should be relatively stable with respect to any uncertainty in the input. That is, if the input is perturbed, the output should vary only slightly, relative to the input perturbation. This is exactly the concept of Lipschitz continuity. If it is possible that the input has been slightly perturbed, the result of the program should not be drastically different.

Lipschitz Robustness Robustness for control systems was investigated by Majumdar and Saha [14]. For programs operating on real numbers, robustness was considered by [11] using the concept of Lipschitz continuity. Additionally, Samanta et al. [12] and Henzinger et al. [13] investigate the use of Lipschitz continuity for proving robustness in the context of transducers.

TODO: incorporate above and below

Verifying the property of Lipschitz continuity is considered by Chaudhuri et al. [10, 11]. The approach taken is to verify continuity of the program as a whole and then verify that each control flow path of the program is piecewise Lipschitz continuous by showing that it is piecewise linear. The continuity verification is performed using a program logic for reasoning about continuity [10]. In order to establish that each control flow path is piecewise linear, an abstract interpretation is applied in which an abstract state, referred to as a robustness matrix, is used to determine each rate of change $\partial x_i^{out} / \partial x_j^{in}$. One advantage of this approach, similar to other relational program logics, is that branches of **if** blocks can be reasoned about independently and then it is verified that the branches are compatible. As this compatibility is transitive, this prevents have to make a

pairwise comparison of each branch as in approaches based on product programs. However, the approach used in [10, 11] is limited to numerical perturbations.

2.3 k -Safety Properties

2-Safety Properties In [6], Terauchi & Aiken introduced the term 2-safety property. A general approach to the verification of 2-safety properties is the creation of a product program [7], a program which interleaves two copies of a given program to create a new program. It [8], Barthe et al. analyzes various relational program logics, as well as different notions of product programs.

TODO: incorporate

A more general approach to verifying robustness properties is the creation of a product program which is then provided as input to a standard program verifier [7]. The product $\mathcal{P}_1 \otimes \mathcal{P}_2$ of two programs $\mathcal{P}_1, \mathcal{P}_2$ is a program which is equivalent to simultaneously executing both \mathcal{P}_1 and \mathcal{P}_2 . Because a single program is created with distinct variables corresponding to each variable of \mathcal{P}_1 and \mathcal{P}_2 , relational properties between \mathcal{P}_1 and \mathcal{P}_2 can be expressed as a standard verification property of the single program $\mathcal{P}_1 \otimes \mathcal{P}_2$. As a result, a 2-safety property about \mathcal{P} , such as robustness, can be established by providing $\mathcal{P} \otimes \mathcal{P}$ to any standard program verifier.

k-Safety Properties Clarkson and Schneider [16] introduced k -safety properties as a generalization of this idea. Sousa and Dillig [17] formulated a verification algorithm in order to automate checking of k -safety properties.

TODO: incorporate

After discussing with the authors of [17], it seems that with a lot of time and effort, their implementation of Descartes could be used to verify many robustness properties.

3 Example Robustness Properties

The uncertainty with respect to which a program is robust can take many different forms. Some common perturbations include

- numerical perturbations [13, 10, 11],
- permutations of arrays and matrices,
- simultaneous permutations of arrays.

3.1 Uniform noise and list sorting

TODO: This section talks about sorting functions being 1-robust. existing work for real numbers works fine for integers.

Next section will mention that continuity work doesn't apply to structural stuff.

As an example, sorting algorithms are 1-Lipschitz. If each element of the input array is perturbed by an amount no more than 1, then each element of the output array will be changed by no more than 1.

3.2 Permutations

TODO: put this discussion in a reasonable place. Break into “invariance” and “respects”

Let F and G be the formulas

$$F : \forall i. 0 \leq i < \text{length}(a_1) \rightarrow a_1[i] = a_2[(i + 1) \% \text{length}(a_1)],$$

$$G : a_1[0] = a_2[1] \wedge a_1[1] = a_2[0] \wedge \forall i. (2 \leq i < \text{length}(a_1) \rightarrow a_1[i] = a_2[i]).$$

Then invariance of a program which takes an array as input can be specified in the language of [17] as

$$\| \text{length}(a_1) = \text{length}(a_2) \wedge (F \vee G) \| \mathcal{P}(a) \| \text{ret}_1 = \text{ret}_2 \|.$$

Sorting Sorting is one example of a procedure which is robust with respect to permuting the input. Even in the face of uncertainty regarding the order of the input array, the result of procedure will not be altered. The `max` function is a special case of this invariance of sorting algorithms under permutation, which is the root of the incorrectness of the attempts to synthesize the `max` function in Section 1.

Searching Consider the function `Find(a, x)` which returns the index of the element `x` in the array `a` or `-1` if `x` is not an element of `a`. Let σ be a permutation and suppose that $a[i] = x$. Then from (1)

$$\sigma a[\sigma(i)] = a[i] = x,$$

so that $\text{Find}(\sigma a, x) = \sigma(i)$. Considering σ as a permutation on nonnegative integers, $\sigma(-1) = -1$ and $\text{Find}(\sigma a, x) = \sigma(i)$ holds even in the case that `x` is not an element of `a`. Thus, `Find` is robust in that perturbing the input array by a permutation σ perturbs the output of `Find` by the same permutation σ .

Adjacency Matrices Big Q: should we include this??

The effect of permuting the rows and columns of an adjacency matrix by the same permutation is simply a relabelling of the vertices. As the labelling of the vertices is arbitrary, a program which takes an adjacency matrix as input should likely satisfy one of two robustness properties. If the program computes a property of the graph, such as the existence of a Hamiltonian cycle, the program should be invariant under such a relabelling. If the program computes a result which depends on the labelling, such as finding an explicit Hamiltonian cycle, the program should satisfy the same property as in the case of searching, that the output should be perturbed by the same permutation as the input.

4 Invariance under permuting lists

TODO: write intro

4.1 Automata

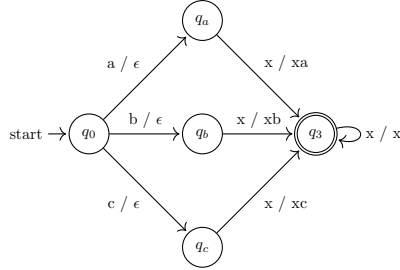
Given a program P which takes a linear data type as input. We may wish to verify that this program is invariant under reordering of the array, in other words, it is invariant under the action of S_n . Denote the output of P on input s by $P(s)$. Then we want to test whether $P(\sigma s) = P(s)$ for all $\sigma \in S_n$.

As a first simplification, observe that it suffices to check our condition on a set of generators of S_n . Concretely, let $\alpha = (1\ 2\ 3\ \dots\ n)$ and let $\beta = (1\ 2)$. Then any $\sigma \in S_n$ can be written as a product of elements of $\{\alpha, \beta\}$, say $\sigma = u_1 u_2 \dots u_m$ with $u_i \in \{\alpha, \beta\}$. So if we know that $P(\alpha s) = P(s)$ and $P(\beta s) = P(s)$, then $P((u_1 u_2 \dots u_m)s) = P((u_2 \dots u_m)s) = \dots = P(u_m s) = P(s)$.

For a fixed n and $\sigma \in S_n$, we can construct a finite state transducer T such that $T(s) = \sigma s$ for $|s| = n$. However, this does not give us any reasonable way of checking that a given finite state machine is invariant under permutation of the input string. Instead, we will construct finite state transducers T_α and T_β such that for any n and any string s with $|s| = n$, we have $T_\alpha(s) = (1\ 2\ \dots\ n)s$ and $T_\beta(s) = (1\ 2)s$. Then, determining whether a FSM M is invariant under permutation of its inputs is equivalent to determining whether

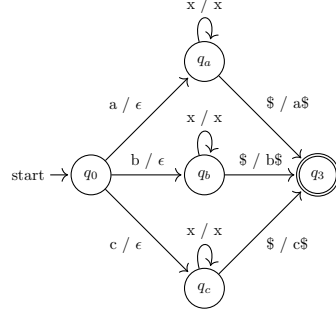
$$L(M \circ T_\alpha) = L(M) = L(M \circ T_\beta)$$

We now construct T_α . For each symbol $a \in \Sigma$, T_α has a transition from its start state s_0 to s_a , while reading in put a and writing ϵ output. For each a and each $b \neq \$$, there is a transition from s_a to s_a , reading b and writing b . Then for each a , there is a transition from s_a to s_1 , reading $\$$ and writing $a\$$.



We now construct T_β . For each symbol $a \in \Sigma$, T_α has a transition from its start state s_0 to s_a , while reading in put a and writing ϵ output. Each s_a has a transition to s_1 while reading input b (for any $b \in \Sigma$) and writing output ba .

Then s_1 simply has transitions to s_1 , reading any $a \in \Sigma$ and writing back a .



4.2 Programs

The previous section indicates a method for verifying robustness of more general programs under permutation of inputs. Given a procedure F taking an array as argument, consider the functions F_α , which first swaps the first two elements of the array, then computes F , and F_β , which moves the first element of the array to the back, then computes F . Then F is invariant under permutation of its input if and only if F is functionally equivalent to F_α and F_β .

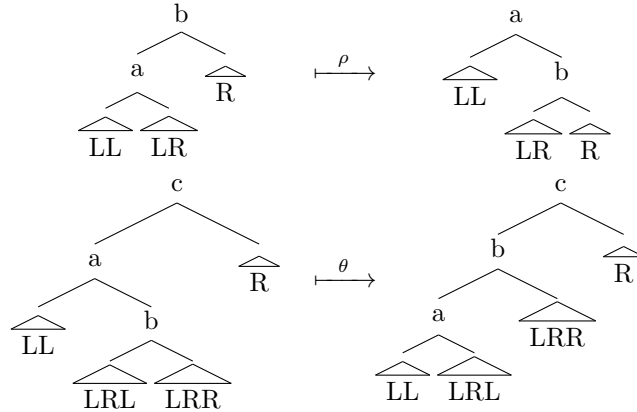
Certainly this method is more efficient than checking invariance under a larger class of permutations. For DFAs, this simplification made the problem of permutation invariance solvable. This indicates that this simplification may aid in our analysis in a profound way. For example, this may be a much easier way of verifying this invariance compared to expressing permutations as a general 2-safety property.

5 Invariance under permuting binary search trees

In the previous section 4, we found a reduction from checking *all* permutations of a list to checking a small set of permutations. It is natural to ask if there are other data types for which we can do this. Binary search trees are one such case, where just like lists, two permutations suffice to generate all equivalent binary search trees (i.e. tree representing equivalent ordered lists).

TODO: Define

We reпре Then define two (partial) operations ρ and θ on binary search trees as follows:



You can verify that ρ and θ preserve $\text{list}(t)$ by observing that the subtrees in the above diagrams remain in the same order before and after applying the transformations. We will show in the next theorem that ρ and θ additionally suffice to generate all transformations of any tree into another tree with an equal underlying list.

Theorem 1. *If a program P on binary search trees is invariant under ρ and θ , then whenever $\text{list}(t_1) = \text{list}(t_2)$, we have $P(t_1) = P(t_2)$.*

Proof. Let t_1 and t_2 satisfy $\text{list}(t_1) = \text{list}(t_2)$. Observe that if P is invariant under ρ and θ , then P is also invariant under ρ^{-1} and θ^{-1} . Thus, it suffices to show that t_1 and t_2 can both be transformed via ρ and θ into another tree t_3 , because then we can transform t_1 into t_3 , then apply the inverse transformation to get to t_2 . We proceed by providing an algorithm for transforming to a t_3 in the following form:

Definition 1. *A tree t is in degenerate list form if either*

1. *t is Nil, or*
2. *t has a Nil left child, and t 's right child is in degenerate list form.*

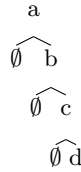


Fig. 1. A tree in degenerate list form

It is clear that if t_1 and t_2 are in degenerate list form, then $\text{list}(t_1) = \text{list}(t_2)$ if and only if $t_1 = t_2$. Thus, if we can show that t_1 and t_2 can both be transformed

into degenerate list form, say t'_1 and t'_2 , then we will have $t'_1 = t'_2$ and we will be done.

Consider the following algorithm for adjusting a tree via ρ and θ , where invariants and assertions are written inside $\{\text{braces}\}$:

```

1: function LISTIFY( $t$ )
2:   while  $t$  has a right child do
3:     apply  $\rho^{-1}$  to  $t$ 
4:    $\{t$  has a null right child $\}$ 
5:   while  $t$  has a left child do  $\triangleright \{t$ 's right child is in degenerate list form $\}$ 
6:     while  $t$ 's left child has a right child do
7:       apply  $\theta$  to  $t$ 
8:      $\{t$ 's left child has a null right child $\}$ 
9:     apply  $\rho$  to  $t$ 
10:   $\{t$  is in degenerate list form $\}$ 

```

To verify that this algorithm terminates with t in degenerate list form, we need only verify that each of the assertions hold and that the loop invariant is inductive.

The invariant on line 4 follows from the negation of the condition of the loop before it. The loop invariant holds from line 4 simply because the null tree is in degenerate list form. The invariant on line 8 also follows from the negation of the condition of the loop before it.

Next we show that the loop invariant is inductive. The inner loop on line 6 does not change t 's right child, so it remain in degenerate list form. Thus, when we apply ρ to t , it's left child has a null right child, and t 's right child is in degenerate list form. As shown in figure 2, this leave $\rho(t)$'s right subtree in degenerate list form. Thus, the invariant is inductive.

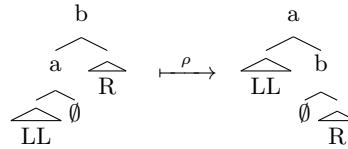


Fig. 2. Applying ρ to a tree as in line 9

Finally, we see that our conclusion on line 10 follows from the loop invariant and the negation of the condition, by the definition of degenerate list form.

6 Invariance with respect to a function

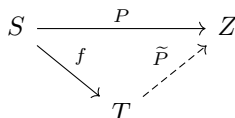
All previous work on these topics share one downside: they require programmers to express their invariants in formal logic. This could potentially be difficult. One class of invariants that can be expressed through codes is *invariance with respect to a function*.

Definition 2. *We say that a program P is invariant with respect to a function f if whenever $f(x) = f(y)$, we have $P(x) = P(y)$.*

For example, if a programmer defines a tree or heap data type with a function list mapping the data type to lists, then nearly every function written on this data type should be invariant with respect to the list function, as these data types are meant to represent the lists perfectly (while allowing for faster algorithms).

We regard this problem as very difficult, because the function f can be expressed with arbitrary code and thus encode more complicated properties than first order logic. One approach comes from the following observation:

Lemma 1. *Let $P : S \rightarrow T$ and $f : T \rightarrow Z$. Then P is invariant with respect to f if and only if there exists a program $\tilde{P} : T \rightarrow Z$ such that $P = \tilde{P} \circ f$*



Proof. If P is invariant with respect to f , then the function

$$\tilde{P}(t) = \begin{cases} P(s), & t = f(s) \text{ for some } s \in S \\ z_0, & \text{otherwise} \end{cases}$$

for any $z_0 \in Z$ is well defined and clearly satisfies $P = \tilde{P} \circ f$.

Conversely, if \tilde{P} exists, then whenever $f(s) = f(s')$, we have $P(s) = \tilde{P}(f(s)) = \tilde{P}(f(s')) = P(s')$, so P is invariant with respect to f .

Now, the key observation is that P and f together give a full functional specification for \tilde{P} . We want to either construct such a \tilde{P} in order to prove our property, or provide the programmer with a counterexample to the property in order to help them debug. Thus, a counterexample guided synthesis loop, similar to that present in [19], seems like a natural candidate for this problem.

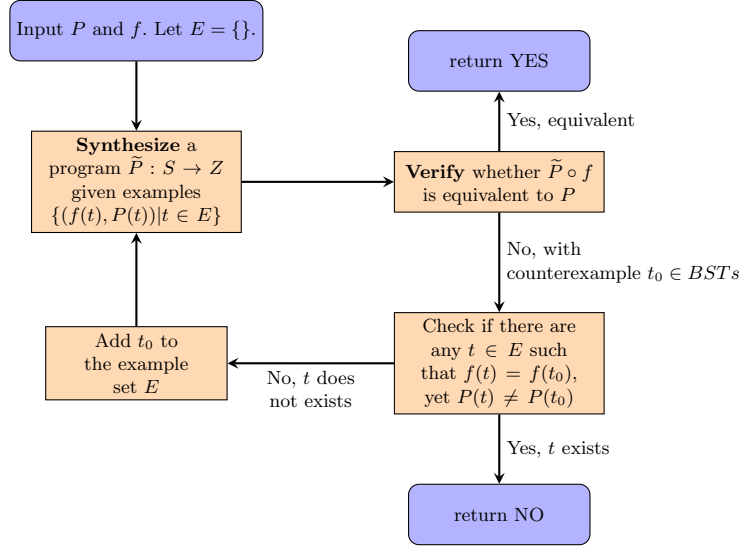


Fig. 3. Decision procedure for verifying invariance with respect to a function

Theorem 2. *The algorithm 3 is sound and relatively complete, relative to a perfect synthesizer and a perfect equivalence-of-programs verifier.*

Proof. In order to prove this formally, we need an even stronger condition on the equivalence-of-programs verifier. Namely, we suppose there is some total order on the elements in T , and that if $\tilde{P} \circ f$ is not equivalent to P , the verifier outputs the smallest counterexample with respect to this ordering.

If \tilde{P} exists, then a perfect synthesizer will eventually synthesize \tilde{P} given some set of examples, say E . Because the verifier gives back counterexamples in a fixed order, a superset of E will eventually be fed into the synthesizer, resulting in \tilde{P} . Then the idealized verifier will output “YES”.

If no \tilde{P} exists, then some pair t, t_0 exist such that $f(t) = f(t_0)$ yet $P(t) \neq P(t_0)$. Because the verifier outputs counterexamples in some order, t and t_0 will eventually be found, and the algorithm will output “YES”.

References

1. Chaudhuri, S., Gulwani, S., Lubliner, R.: Continuity and robustness of programs. *Commun. ACM* **55**(8) (August 2012) 107–115
2. Benton, N.: Simple relational correctness proofs for static analyses and program transformations. In: *Proceedings of the 31st ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages*. POPL ’04, New York, NY, USA, ACM (2004) 14–25
3. Barthe, G., Grégoire, B., Zanella Béguelin, S.: Formal certification of code-based cryptographic proofs. In: *Proceedings of the 36th Annual ACM SIGPLAN-*

- SIGACT Symposium on Principles of Programming Languages. POPL '09, New York, NY, USA, ACM (2009) 90–101
4. Barthe, G., Köpf, B., Olmedo, F., Zanella-Béguelin, S.: Probabilistic relational reasoning for differential privacy. *ACM Trans. Program. Lang. Syst.* **35**(3) (November 2013) 9:1–9:49
 5. Barthe, G., D’Argenio, P., Rezk, T.: Secure information flow by self-composition. In: *Proceedings of the Computer Security Foundations Workshop*. Volume 17. (2004) 100–114
 6. Terauchi, T., Aiken, A.: Secure information flow as a safety problem. In Hankin, C., Siveroni, I., eds.: *Static Analysis: 12th International Symposium, SAS 2005*, London, UK, September 7-9, 2005. *Proceedings*, Berlin, Heidelberg, Springer Berlin Heidelberg (2005) 352–367
 7. Barthe, G., Crespo, J.M., Kunz, C.: Product programs and relational program logics. *Journal of Logical and Algebraic Methods in Programming* **85**(5, Part 2) (2016) 847 – 859 Articles dedicated to Prof. J. N. Oliveira on the occasion of his 60th birthday.
 8. Barthe, G., Crespo, J.M., Kunz, C.: Relational verification using product programs. In Butler, M., Schulte, W., eds.: *FM 2011: Formal Methods: 17th International Symposium on Formal Methods*, Limerick, Ireland, June 20-24, 2011. *Proceedings*, Berlin, Heidelberg, Springer Berlin Heidelberg (2011) 200–214
 9. Hamlet, D.: Continuity in software systems. In: *Proceedings of the 2002 ACM SIGSOFT International Symposium on Software Testing and Analysis. ISSTA '02*, New York, NY, USA, ACM (2002) 196–200
 10. Chaudhuri, S., Gulwani, S., Lubliner, R.: Continuity analysis of programs. In: *Proceedings of the 37th Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages. POPL '10*, New York, NY, USA, ACM (2010) 57–70
 11. Chaudhuri, S., Gulwani, S., Lubliner, R., Navidpour, S.: Proving programs robust. In: *Proceedings of the 19th ACM SIGSOFT Symposium and the 13th European Conference on Foundations of Software Engineering. ESEC/FSE '11*, New York, NY, USA, ACM (2011) 102–112
 12. Samanta, R., Deshmukh, J.V., Chaudhuri, S.: Robustness analysis of string transducers. In Van Hung, D., Ogawa, M., eds.: *Automated Technology for Verification and Analysis: 11th International Symposium, ATVA 2013*, Hanoi, Vietnam, October 15-18, 2013. *Proceedings*, Cham, Springer International Publishing (2013) 427–441
 13. Henzinger, T.A., Otop, J., Samanta, R.: Lipschitz Robustness of Finite-state Transducers. In Raman, V., Suresh, S.P., eds.: *34th International Conference on Foundation of Software Technology and Theoretical Computer Science (FSTTCS 2014)*. Volume 29 of *Leibniz International Proceedings in Informatics (LIPIcs)*., Dagstuhl, Germany, Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik (2014) 431–443
 14. Majumdar, R., Saha, I.: Symbolic robustness analysis. In: *2009 30th IEEE Real-Time Systems Symposium*. (Dec 2009) 355–363
 15. Samanta, R., Deshmukh, J.V., Chaudhuri, S.: Robustness analysis of networked systems. In Giacobazzi, R., Berdine, J., Mastroeni, I., eds.: *Verification, Model Checking, and Abstract Interpretation: 14th International Conference, VMCAI 2013*, Rome, Italy, January 20-22, 2013. *Proceedings*, Berlin, Heidelberg, Springer Berlin Heidelberg (2013) 229–247
 16. Clarkson, M.R., Schneider, F.B.: Hyperproperties. In: *2008 21st IEEE Computer Security Foundations Symposium*. (June 2008) 51–65

17. Sousa, M., Dillig, I.: Cartesian hoare logic for verifying k-safety properties. In: Proceedings of the 37th ACM SIGPLAN Conference on Programming Language Design and Implementation. PLDI '16, New York, NY, USA, ACM (2016) 57–69
18. Dummit, D.S., Foote, R.M.: Abstract Algebra. 3rd edn. John Wiley & Sons (2004)
19. Solar-Lezama, A., Tancau, L., Bodik, R., Seshia, S., Saraswat, V.: Combinatorial sketching for finite programs. In: Proceedings of the 12th International Conference on Architectural Support for Programming Languages and Operating Systems. ASPLOS XII, New York, NY, USA, ACM (2006) 404–415

A Preliminaries

A.1 Lipschitz Robustness

Lipschitz continuity of a mathematical function f is the property that there exists a real number $K \in \mathbb{R}^+$ so that for all x_1, x_2 , $d(f(x_1), f(x_2)) \leq K d(x_1, x_2)$. Similarly, a program \mathcal{P} is robust at a state σ with respect to the input variable x_{in} and output variable x_{out} if for all $\varepsilon \in \mathbb{R}^+$, if σ' satisfies

1. $d(\sigma(x_{in}), \sigma'(x_{in})) < \varepsilon$ and
2. for all $y \neq x$, $\sigma(y) = \sigma'(y)$,

then $d(\mathcal{P}(\sigma)(x_{out}), \mathcal{P}(\sigma')(x_{out})) < K\varepsilon$, where K is a real-valued function of the size of x_{in} [11].

A.2 Permutations

A permutation is a bijection from a set Ω to itself [18]. When Ω is a finite set, a permutation can be specified using two-line notation by placing the elements of Ω on one line, and their images under σ beneath them:

$$\sigma: \begin{array}{cccccc} 0 & 1 & 2 & 3 & 4 & 5 \\ 5 & 2 & 3 & 1 & 4 & 0 \end{array} \quad \tau: \begin{array}{ccc} a & b & c \\ a & c & b \end{array}$$

Similar to two-line notation, a permutation on $\{0, \dots, N-1\}$ can be encoded in an array a of length N by placing the image of i in $a[i]$. The example σ above is encoded in an array as $\sigma = [5, 2, 3, 1, 4, 0]$, so that $\sigma(i) = \sigma[i]$.

Action on an Array Given a permutation σ , σ can *act* on an array a as follows. Let $a = [37, 25, 19, 49, 81, 21]$ and σ be as above. Then $\sigma(3) = \sigma[3] = 1$ and applying σ to a results in the element of a at index 3, 49, being moved to index 1 in σa . That is, $\sigma a[1] = 49$ and $a[3] = \sigma a[1] = \sigma a[\sigma[3]]$. More generally,

$$\sigma a[\sigma[i]] = a[i]. \tag{1}$$

Generating Permutations Information on permutation groups can be found in §1.3 of [18].

The permutations on N elements, S_N , can all be expressed as compositions of the permutations [Exercises 3-4, §3.5, 18]

$$\begin{array}{cccccc} 0 & 1 & 2 & \cdots & N-1 \\ 1 & 0 & 2 & \cdots & N-1 \end{array} \quad \text{and} \quad \begin{array}{cccccc} 0 & 1 & 2 & \cdots & N-1 \\ 1 & 2 & 3 & \cdots & 0 \end{array},$$

where the first permutation is a transposition of 0 and 1 and the second permutation is a rotation of all of the elements by one position.

A.3 Automata

We consider two classes of automata: finite state machines (FSM) and finite state transducers (FST). All automata we consider are deterministic. Given an automata A , we denote by $A(s)$ the output of A on input s . If A is a FSM, we represent “accept” by 1 and “reject” by 0. Otherwise, A is a finite state transducer (FST), and $A(s)$ is a string. For a FSM A , let $L(A)$ denote the set of strings accepted by A . (We do not define the language of a FST).

Recall that for any FSMs A and B , the problem of determining whether $L(A) = L(B)$ is decidable. We can compose a FST T and a FSM A , denoted $A \circ T$, to get another FSM such that $(A \circ T)(s) = A(T(s))$. Furthermore, $A \circ T$ can be constructed with $|S_A||S_T|$ states, where S_A and S_T are the states of A and T . We will assume that all input strings are terminated by a special end-of-input character \$.