Ordered t-way Combinations for Testing State-based Systems

D. Richard Kuhn  
M S Raunak  
*Computer Security Division*  
*Information Technology Laboratory*

Raghu N. Kacker  
*Applied and Computational Mathematics Division*  
*Information Technology Laboratory*

June 13, 2022

This publication is available free of charge from:  
https://doi.org/10.6028/NIST.CSWP.26
Abstract

Fault detection often depends on the specific order of inputs that establish states which eventually lead to a failure. However, beyond basic structural coverage metrics, it is often difficult to determine if code has been exercised sufficiently to ensure confidence in its functions. Measures are needed to ensure that relevant combinations of input values have been tested with adequate diversity of ordering to ensure correct operation. Combinatorial testing and combinatorial coverage measures have been applied to many types of applications but have some deficiencies for verifying and testing state-based systems where the response depends on both input values and the current system state. In such systems, internal states change as input values are processed. Examples include network protocols, which may be in listening, partial connection, full connection, disconnected, and many other states depending on the values of packet fields and the order of packets received. Similarly, merchant account balances in credit card systems change continuously as transactions are processed. This publication introduces a notion of ordered t-way combinations, proves a result regarding the construction of adequate blocks of test inputs, and discusses the application of the results to verify and test state-based systems.

Keywords

combinatorial coverage; combinatorial methods; combinatorial testing; software testing; structural coverage; test coverage.

Disclaimer

Any mention of commercial products or reference to commercial organizations is for information only; it does not imply recommendation or endorsement by NIST, nor does it imply that the products mentioned are necessarily the best available for the purpose.

Additional Information

For additional information on NIST’s Cybersecurity programs, projects and publications, visit the Computer Security Resource Center. Information on other efforts at NIST and in the Information Technology Laboratory (ITL) is also available.

Submit comments on this publication to: cfdwp@nist.gov

National Institute of Standards and Technology
Attn: Computer Security Division, Information Technology Laboratory
100 Bureau Drive (Mail Stop 8930) Gaithersburg, MD 20899-8930
1 Introduction

Vulnerability and fault detection often depend on the specific order of inputs that establish states which eventually lead to a failure. That is, many software processes are not deterministic functions where an event eventually lead to a failure. That is, many software specific order of inputs that establish states which is by using some sort of structural coverage metrics, evaluating how thoroughly software has been tested effective software testing. The common practice for encountered in practice is critical to any form of sufficiently representative of what will be Ensuring that inputs and system states in testing are sufficiently representative of what will be encountered in practice is critical to any form of effective software testing. The common practice for evaluating how thoroughly software has been tested is by using some sort of structural coverage metrics, such as statement coverage, branch coverage, or MC/DC coverage [13] . Test cases selected using only structural coverage criteria are often not very effective as they are not designed to include corner cases with specific combinations of input values that may cause a failure. Looking beyond these commonly used metrics, it is often difficult to determine if the code has been adequately tested and even more difficult to ensure that a sufficient diversity of inputs has been achieved. This is particularly true of assertion-based testing or runtime verification, where program states and properties are monitored to verify correct processing. For runtime assertions to discover bugs, the software needs to be exercised with a set of values in a particular order that leads to the failure. Consequently, for strong software assurance, measures are needed to verify that combinations of input values and combinations of input orders in a test suite are sufficient.

Combinatorial coverage measures provide an effective method for quantifying the thoroughness of test input values [26] . A number of measures have been defined for the coverage of (static) input value combinations. For example, with four binary variables, there are a total of \(2^4 \times C(4, 2) = 24\) possible settings of the four variables taken two at a time. If a test set includes tests that cover 19 of the 24, the simple combinatorial coverage is \(19/24 = 0.79\). These measures quantify the degree to which input values cover the potential space of parameter value combinations without regard for the order in which these inputs occur in a test set or in normal operations. However, if a system state is affected or determined by the order of inputs, even thorough coverage of the input space may not detect some failure conditions. Thus, it is desirable to supplement measures of input space coverage with measures of the input value combination ordering.

2 Related Work

Combinatorial aspects of input ordering have been studied in the context of event sequences. Sequence covering arrays (SCAs) were introduced in [1] and [2] and further developed in [3], [4], [5], [7], [8] [9], [10], and [11].

Definition. A sequence covering array [2], SCA(N, S, t) is an \(N \times S\) matrix where entries are from a finite set \(S\) of \(s\) symbols, such that every \(t\)-way permutation of symbols from \(S\) occurs in at least one row. The \(t\) symbols in the permutation are not required to be adjacent.

For example, Fig. 1 shows an event sequence \(a \rightarrow b \rightarrow c\) in test 1 and an event sequence of \(d \rightarrow e \rightarrow a\) in test 3, where \(x \rightarrow y\) denotes \(x\) is eventually followed by \(y\), with possible interleaving. Note that the event sequence array has sequences of events in each row. Event sequences are made up of a value in a column followed by values in columns to the right.

<table>
<thead>
<tr>
<th>Test</th>
<th>p0</th>
<th>p1</th>
<th>p2</th>
<th>p3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>a</td>
<td>d</td>
<td>b</td>
<td>c</td>
</tr>
<tr>
<td>2</td>
<td>b</td>
<td>a</td>
<td>c</td>
<td>d</td>
</tr>
<tr>
<td>3</td>
<td>b</td>
<td>d</td>
<td>c</td>
<td>a</td>
</tr>
<tr>
<td>4</td>
<td>c</td>
<td>a</td>
<td>b</td>
<td>d</td>
</tr>
<tr>
<td>5</td>
<td>c</td>
<td>a</td>
<td>b</td>
<td>a</td>
</tr>
<tr>
<td>6</td>
<td>d</td>
<td>a</td>
<td>c</td>
<td>b</td>
</tr>
</tbody>
</table>

Fig. 1. Event sequence
Combinatorial testing with constraints on the order in which values and combinations are applied in tests was analyzed extensively in [5]. Extended covering arrays that consider the sequence of values in each test were defined in [6]. Combination sequences were studied in [20], combining configuration and the order of combinations while also considering constraints. Another structure defined as a sequence covering array of $t$-way combinations has been termed a multi-valued sequence covering array, introduced in [25].

3 Ordered Combination Covering

A combination order is different from an event sequence. As noted in the previous section, an event sequence is a possibly interleaved sequence of symbols in a single row of a test array. A combination order, as defined below, is across multiple rows, given in the order in which tests will be executed. A $t$-way permutation of symbols is referred to as a $t$-way order, which will be called a $t$-order for brevity. The $t$ events in the order may be interleaved with others (i.e., the order $a \rightarrow \text{b} \rightarrow \text{c}$ covers three $2$-event orders: $a$ followed by $\text{b}$ and $\text{b}$ followed by $\text{c}$, and $a$ followed by $\text{c}$). Denoting event $a$ eventually followed by event $\text{b}$, possibly with other events interleaved, is written as $a{}*{}\rightarrow{}b$.

Consider the notion of an $s$-order of $t$-way combinations of the input parameters as a series of rows of test data that contain a particular set of $t$-way combinations in a specified order, with possibly interleaved rows.

**Definition 1.** The notation $a{}*{}\rightarrow{}b$ denotes the presence of combination $a$ eventually followed by combination $b$, possibly with other rows interleaved. A combination order $c_1{}*{}\rightarrow{}c_2{}*{}\rightarrow{}...*{}\rightarrow{}c_s$ of $s$ sequence of $t$-way combinations, abbreviated $s$-order, is a set of $t$-way combinations in $s$ rows. Each $c_i$ is a $t$-way combination of parameter values.

**Example.** Fig. 2 shows combination order $p_0 p_1 = \text{ad} \rightarrow p_0 p_3 = \text{bd} \rightarrow p_1 p_3 = \text{ab}$, which is a 3-order of 2-way combinations. Thus, the term ordered combinations refers to combinations in a row followed by combinations in rows below.

When all $s$-orders of $t$-way combinations of the input parameters have been covered, it is referred to as an ordered combination cover (OCC). For the OCCs, the combination orders are treated across rows (i.e., a combination in a row followed by combinations in rows below). A $t$-way combination occurs in some row and is eventually followed by other $t$-way combinations in other rows. For three Boolean parameters $a, b, c$ in Fig. 3, $a = 00$ is followed by $ab = 01, ac = 01$ and $bc = 01$ (ac=01 and bc=01 are also followed by this group).

<table>
<thead>
<tr>
<th>Test</th>
<th>$p_0$</th>
<th>$p_1$</th>
<th>$p_2$</th>
<th>$p_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\text{a}$</td>
<td>$\text{d}$</td>
<td>$\text{b}$</td>
<td>$\text{c}$</td>
</tr>
<tr>
<td>2</td>
<td>$\text{b}$</td>
<td>$\text{a}$</td>
<td>$\text{c}$</td>
<td>$\text{d}$</td>
</tr>
<tr>
<td>3</td>
<td>$\text{b}$</td>
<td>$\text{d}$</td>
<td>$\text{c}$</td>
<td>$\text{a}$</td>
</tr>
<tr>
<td>4</td>
<td>$\text{c}$</td>
<td>$\text{a}$</td>
<td>$\text{b}$</td>
<td>$\text{d}$</td>
</tr>
<tr>
<td>5</td>
<td>$\text{c}$</td>
<td>$\text{d}$</td>
<td>$\text{b}$</td>
<td>$\text{a}$</td>
</tr>
<tr>
<td>6</td>
<td>$\text{d}$</td>
<td>$\text{c}$</td>
<td>$\text{a}$</td>
<td>$\text{b}$</td>
</tr>
</tbody>
</table>

Fig 2. Ordered combinations

Combinations of parameter values such as this can be significant in protocol verification and testing, such as with combinations of values in each packet affecting the state where later packets produce different responses depending on the state that resulted from previous packet orders.

**Definition 2.** An ordered combination cover, OCC($N, s, t, p, v$), covers all $s$-orders of $t$-way combinations of the $v$ values of $p$ parameters, where $t$ is the number of parameters in combinations and $s$ is the number of combinations in an ordered series. Permutations of parameter value combinations may appear multiple times in a combination order. For example, a particular 2-order of 2-way combinations may be $(p_1 p_2 = 01) \rightarrow (p_2 p_4 = 11)$.

The utility of combination order covering can be illustrated with an example. Consider the covering array in Fig. 4, which includes all 2-way combinations of four Boolean variables.
Suppose these six tests are applied to the system modeled with a finite state machine diagram in Fig. 5, and tests are run in the order 1..6. If condition A = “p1 ∧ p2,” and condition B = “p1 ∧ ¬p2”, then the error in state 2 will not be discovered. The system returns to state 0 for tests 1 through 3, then enters state 1 with test 4 and moves to state 3 with test 5. Because the test array does not include the ordered combinations p1, p2 = 11 *→ p1, p2 = 10, the error is not exposed. However, if the tests are run in the order [1, 2, 4, 3, 5, 6], then the error in state 2 will be discovered because the third test (row 4 in Fig. 4) leads to state 1, the fourth test (row 3) causes condition B to evaluate to true, and the system enters state 2, exposing the error.

When tests are executed in sequence with each individual t-way combination considered an event, an order of t-way combinations containing s combinations input in sequence with possible interleaving is an s-sequence of t-way combinations. For example, a 2-sequence of 3-way combinations could be

$$abd = 001 \rightarrow bcd = 100,$$

and a 3-sequence of 2-way combinations could be

$$be = 01 \rightarrow ad = 11 \rightarrow bc = 10.$$
reduced to ensuring coverage of combination sequence coverage can be and resource intensive. Fortunately, the problem of combination sequence coverage can be inefficient in expression (1). Consequently, measuring realistic testing problems as a result of the exponents sequences to be covered will become very large with.

As seen in Fig. 6, the numbers of combination covering arrays in Testing

3.1 Combination Sequence Covering Arrays in Testing

As seen in Fig. 6, the numbers of combination sequences to be covered will become very large with realistic testing problems as a result of the exponents in expression (1). Consequently, measuring combination sequence coverage can be inefficient and resource intensive. Fortunately, the problem of ensuring combination sequence coverage can be reduced to ensuring coverage of t-way covering arrays, as shown in the following proof. Checking that a test array is a covering array can be done efficiently, making it practical to ensure s-sequences of t-way combinations in large-scale testing.

Theorem: A test set covers s-orders of t-way combinations if and only if it includes an ordered series containing a total of s covering arrays, each of strength t.

Proof: From Definition 2, a sufficient process for generating an s-order t-way OCC is to concatenate s covering arrays of strength t, as shown below in Fig. 7. Because a covering array includes every t-way combination, any order of at least s combinations will occur by taking the rows of s covering arrays from CA1, CA2, ..., CA_s, where CA_s are t-way covering arrays. Clearly, for any s-order of s combinations, c_1 → c_2 → ... → c_s, c_1 must be present in CA_1, c_2 in CA_2, etc. because they cover all t-way combinations by definition, giving the required order.

Fig.6. Tests for four parameters, OCC(20,2,4,2)

That is, the test array includes every solution of (p_i p_x = v_1 v_2) → (p_i' p_x = v_1' v_2), of which there are \((C(4,2) \times 2^2)^2 = 576\) instances. For example, each of the four possible settings of p_i p_x is followed by each of the four possible settings of p_i p_x, somewhere in the table (distinguished by color). That is, (p_i p_x = 11) in line 1 is followed by (p_i p_x = 01), (p_i p_x = 01), (p_i p_x = 11), (p_i p_x = 10), (p_i p_x = 10) in lines 3, 4, 5, and 7, respectively, highlighted in yellow (also for (p_i p_x = 00)). Sequences for (p_i p_x = 01) are shown in green, plus line 14, which provides a (p_i p_x = 00) for both (p_i p_x = 01) and (p_i p_x = 10). Sequences for (p_i p_x = 10) are highlighted in blue.

### 3.1 Combination Sequence Covering Arrays in Testing

To show necessity, consider a series of rows in a test array. There must be at least one combination order that can only exist if the test array can be divided into subarrays, each of which is a covering array. Each row covers some number of t-way combinations. For each row, add the combinations covered to a set, and continue adding non-covered combinations from each successive row. Eventually, a row will be reached that covers the last remaining previously uncovered combinations. Label these previously uncovered combinations A_1 ... A_k and the row containing these combinations as row i. A_1 ... A_k do not occur in any row prior to row i. The subarray of rows from the first row to and including row i forms a covering array that will be labeled CA_i. With the inclusion of row i, CA_i includes all t-way combinations, so it is a t-way covering array.
At row $i+1$, start a new set of combinations covered in rows $i+1$ and following rows. Continue adding combinations covered in each successive row until a row is reached that covers the last remaining previously uncovered combinations after row $i+1$. Label these previously uncovered combinations $\mathcal{B}_1 \ldots \mathcal{B}_x$ and the row containing these combinations as row $x$. $\mathcal{B}_1 \ldots \mathcal{B}_x$ do not occur in any row after row $i$ and prior to row $x$. The subarray beginning with row $i+1$ and ending with row $x$ forms a covering array that will be labeled CA$_2$. CA$_2$ includes all $t$-way combinations, with the inclusion of row $x$, so it is a $t$-way covering array. Any combination in $A_1 \ldots A_2 \ldots A_k$ must be followed by any $t$-way combination in some row of CA$_2$.

Note that any 2-order $c_1 \rightarrow c_2$ where $c_1$ is one of $A_i$ and $c_2$ is one of $\mathcal{B}_x$ could not have been covered until row $x$ of CA$_2$ because the $\mathcal{B}_x$ tuples are those that had not been covered in CA$_2$ before row $x$ (and after row $i$). Assume that $c_1 \rightarrow c_2$ is covered before row $x$ in the combined array CA$_1$ || CA$_2$. Since $c_2$ is not in subarray CA$_2$ before row $x$, it must be in subarray CA$_1$. However, $c_1$ is in the last row of CA$_1$, so $c_2$ must be in a row of CA$_1$ following the last row of CA$_1$, which is a contradiction.

Therefore $c_1 \rightarrow c_2$ can be covered only if CA$_1$ and CA$_2$ are covering arrays. Continuing in this manner shows that orders of $s$ combinations of strength $t$ can be covered only if the subarrays of the set of test rows form $s$ covering arrays (end proof).

**Example.** Fig. 8 shows the concatenation of two 2-way covering arrays for four binary parameters. Any 2-order of 2-way combinations occurs somewhere in the rows of Fig. 6. For example, $(p_1p_2 = 10)$ (row 3) is followed by $p_3p_4 = 00, 01, 10, 11$ in rows 5, 6, 9, and 4, respectively. If row 12 is removed, there must be at least one combination order $c_1 \rightarrow c_2$ where $c_2 = (p_3p_4 = 11)$ that is not covered because $(p_3p_4 = 11)$ is covered only in the last row of CA$_1$ and CA$_2$ (row 12). Removing row 12 would result in losing $(p_3p_4 = 11) \rightarrow (p_3p_4 = 11)$. Similarly, $(p_3p_4 = 10)$ is covered only in the third-to-last row (row 10) of CA$_1$ and CA$_2$, so there must be at least one combination order $c_1 \rightarrow c_2$ where $c_2 = (p_3p_4 = 10)$ that is not covered if row 10 is removed. Removing row 10 would result in losing $(p_3p_4 = 11) \rightarrow (p_3p_4 = 10), (p_3p_4 = 00) \rightarrow (p_3p_4 = 10)$, and others.

The practical utility of this result is that it shows one can efficiently produce tests that cover all orders of $t$-way combinations up to any necessary order length by concatenating $t$-way covering arrays. It also shows that the minimum size of the OCC is determined by the minimum size of the relevant $t$-way covering arrays. From a testing perspective, producing full coverage of $t$-way combinations in $s$ length orders makes it possible to detect faults that are only detectable when a system is in a particular state that can only be reached by an order of input combinations.

<table>
<thead>
<tr>
<th></th>
<th>p1</th>
<th>p2</th>
<th>p3</th>
<th>p4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Fig. 8. Two concatenated covering arrays.

This result can also be useful for runtime verification, assertion monitoring, and other methods that rely on checking program properties and states as code is executed. If inputs are monitored and recorded, then it is possible to verify whether a covering array series of desired length has been applied in testing. The system should run long enough to enter all major states and allow detection of errors that occur only in particular states. The use of covering arrays gives stronger assurance that relevant states have been reached, as program states depend on the order of inputs, and the coverage of input value combinations can be measured.

**3.2 Combination Order Coverage Measurement**

A combination order tool for OCCs, Corder, has been developed, allowing for the order coverage of any test set to be measured. It may also be used in generating OCCs using random test generation, measuring coverage, and extending the test array until sufficient coverage is achieved.
In its current form, the tool assumes that all single values of input variables have been included in the input test array and computes $t$-way coverage for $t = 2..4$ in the same manner as the CCM tool for combination coverage [26]. This is referred to in the output report as static coverage and measures the coverage of combinations in each row of the array where any $t$-way covering array will have 100% coverage of $t$-way combinations. A second output provides coverage, referred to as dynamic, of $s$-orders of $t$-way combinations in the test array.

For example, the test array in Fig. 9 (a) shows 12 tests with four binary parameters or variables. If these are executed in order, the first test includes $C(4,2) = 6$ events defined as 2-way combinations: $ab = 00$, $ac = 01$, $ad = 00$, $bc = 01$, $bd = 00$, and $cd = 10$. For 2-orders containing $ab$, there are four possible settings of $ab$. Each of these may followed by value combinations of $ab$, $ac$, $ad$, $bc$, $bd$, and $cd$. Completely covering all 2-way 2-orders (i.e., orders of length 2 of 2-way combinations) would produce $4 \times C(4,2) \times 4 \times C(4,2)$ orders. One can measure the degree to which these orders are covered and output any missing orders, as shown in Fig. 4(b). Note that $ab = 11$ is followed by $cd = 01$ and $cd = 10$, but $cd = 00$ and $cd = 11$ do not follow $ab = 11$ in the test series, as shown in Fig. 5(c), which shows \langle parameter numbers\rangle : <order> \rightarrow <parameter numbers> : <order> for 2-way orders.

![Fig. 9. Missing combination orders](image)

Fig. 8 illustrates the output of the Corder tool for the matrix shown in Fig. 4. Basic static coverage measures are shown in the top half of the results to provide an overview of input space coverage. For more detailed data on input space coverage, the CCM tool measuring combination coverage can be used [26]. An example is shown in Fig. 11 for a larger, more realistic array of 10 parameters with 6 values each in a file of 263 tests.

![Fig. 11. 10-parameter combination order cover measurement](image)

It is important to understand the difference between static and dynamic coverage as defined here. Essentially, static coverage is based on the presence or absence of $t$-way settings of the input variables, and dynamic coverage refers to the coverage of possible orders of these combinations.
Static coverage corresponds to combination coverage measures defined in [26] and other publications. In this case, there are four variables with two values each, so for single-variable coverage (t=1), there are 4x2 = 8 possible settings. The tool assumes that all single values of variables are included in the input file. For 2-way combinations, there are C(4,2) = 6 possible combinations, each of which has 2x2 = 4 possible settings, so the number of combinations to be covered is 24, all of which are covered at least once. Similarly, for 3-way combinations there are C(4,3) = 4 combinations, which may have 2x2x2 = 8 settings for a total of 32 possible settings to be covered. Of the possible settings, there are 22 covered (for example, abc = 000 is missing, as is acd = 000) for a coverage figure of .6875.

For dynamic coverage, the interaction strength (level of t) for the combinations included in orders and the number of combinations in an order need to be specified.

4 Conclusions

This report considers practical methods for testing complex orders of all t-way combinations up to some specified level of t. It is shown that the test set covers s-orders of t-way combinations if and only if it includes an ordered series of s covering arrays of strength s. This result can efficiently produce tests that cover all orders of t-way combinations up to any necessary order length by concatenating t-way covering arrays. The notion of ordered combination covers may be applied in runtime verification, assertion monitoring, and other verification and test methods that rely on checking program properties and states as code is executed.

5 References

[10] Binder RV. Optimal scheduling for combinatorial software testing and design of experiments. In2018 IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW) 2018 Apr 9 (pp. 295-301). IEEE.
[11] Z. Ratliff, C SCM Manual. NIST. (Note: this coverage measurement tool was originally named C SCM. It has been renamed ESeq to more easily distinguish between the two forms of sequence coverage.)
[19] Rabbi K, Mamun Q, Islam MR. A Novel Swarm Intelligence Based Sequence Generator. Intl Conference on
Applications and Techniques in Cyber Security and Intelligence 2017 Jun 16 (pp. 238-246). Edizioni della Normale, Cham.

Extensions


Disclaimer: Products may be identified in this document, but identification does not imply recommendation or endorsement by NIST, nor that the products identified are necessarily the best available for the purpose.