

Assignment 3 – Design of Experiments

Name – Prasad Aggarwal

ASU id – paggar10

Research and Identifying the DOE tool to Use

Tools researched for Design of Experiments:

1. CAgen
2. AllPairs
3. PICT
4. ACTS

Shortlisted tool for DOE – **CAgen**

Reasons for choosing CAgen as the preferred tool:

- User-Friendly Interface
 - CAgen has a very intuitive graphical user interface that makes it easy to use. You can simply upload the ACTS file, configure parameters, generate tests, and export results through the interactive web UI without any complex setup. Tools like AllPairs and PICT require command line usage which is harder.
- Flexible Coverage Options
 - CAgen supports different coverage criteria like pairwise, 3-wise, 4-wise etc. and you can choose what fits your testing needs. Other tools like AllPairs only offer pairwise while ACTS has limited options. CAgen gives you more flexibility.
- Advanced Test Optimization
 - CAgen has in-built algorithms that minimize the test suite size while maximizing coverage for your configurations. This optimization ensures you get a compact and effective set of tests. The open-source tools don't have the same level of optimization.
- Seamless Integration
 - CAgen allows exporting tests in CSV, Java, XML, etc. which makes integration with testing frameworks straightforward. ACTS has custom formats while AllPairs requires scripting to integrate tests. CAgen provides seamless integration options.

In summary, CAgent provides a top-class user experience, flexibility, optimization, and integration capabilities that give it an edge over open-source DOE tools for this testing assignment. The ease-of-use and effectiveness made it the best choice for generating combinatorial tests via a web-based interface.

DOE tool:

Below is my ACTS configuration file I uploaded on CAgent:

[System]

Name: assignment3

[Parameter]

Type of Phone(enum): iPhone 14,iPhone 13,Galaxy Z,Huawei Mate,Google Pixel 7

Authentication(enum): Fingerprint,Face recognition,Text password

Connectivity(enum): Wireless,3G,4G LTE,5G

Memory(enum): 128GB,256GB,512GB,1TB

Battery(enum): <20%,20-39%,40-59%,60-79%,80-100%

[Constraint]

Total test cases v/s Pairwise methodology:

Pairwise (or all-pairs) testing is a combinatorial testing technique where the objective is to cover all possible discrete pairs of parameter values in the test cases. To generate optimal pairwise test suites, there are two key steps:

- **Parameter Ordering**
 - The parameters are arranged in descending order of their number of values. This ordering helps minimize the total number of test cases required. In our case, the parameters are ordered as:
 - Type of Phone (5 values)
 - Battery (5 values)
 - Authentication (3 values)
 - Connectivity (4 values)
 - Memory (4 values)
- **Pairwise Calculation**
 - With parameters ordered, the number of test cases needed is calculated as:
 - Test cases = \prod (number of values for each parameter). Here, the total test cases required = $5 \times 3 \times 4 \times 4 \times 5 = 1200$

However, combinatorial test generation tools like CAgen implement optimization algorithms that produce the minimal set of tests for pairwise coverage. In this case, CAgen generated only 25 test cases that cover all possible pairwise combinations between the parameters and their values in a highly optimized manner. In summary, ordering parameters and calculating pairwise combinations provides a systematic methodology to generate optimized test suites for effective pairwise testing with minimal tests. This improves testing efficiency and defect detection capability.

Screenshot of test cases generated by CAgen (DOE tool):

v1.2
CAgen
assignment3

Workspaces
Input Parameter Model
Generate
Help
Release Notes
About
Downloads

Array Generation

Algorithm: FIPOG
t: 2
λ: 1
Generate

TEST SET
t=2 25 rows
Randomize Don't-Care Values
Show model values
Export...

| Type of Phone | Authentication | Connectivity | Memory | Battery |
|---------------|------------------|--------------|--------|---------|
| iPhone 14 | Fingerprint | Wireless | 128GB | <20% |
| iPhone 14 | Face recognition | 3G | 256GB | 20-39% |
| iPhone 14 | Text password | 4G LTE | 512GB | 40-59% |
| iPhone 14 | Fingerprint | 5G | 1TB | 60-79% |
| iPhone 14 | Text password | Wireless | 256GB | 80-100% |
| iPhone 13 | Fingerprint | 3G | 512GB | <20% |
| iPhone 13 | Face recognition | 4G LTE | 1TB | 20-39% |
| iPhone 13 | Text password | 5G | 128GB | 40-59% |
| iPhone 13 | Face recognition | Wireless | 256GB | 60-79% |
| iPhone 13 | Text password | 3G | 1TB | 80-100% |
| Galaxy Z | Fingerprint | 4G LTE | 256GB | <20% |
| Galaxy Z | Face recognition | 5G | 512GB | 20-39% |

Complete CSV file of the generated test cases from CAgen:

assignment3-t2

| Type of Phone | Authentication | Connectivity | Memory | Battery |
|----------------|------------------|--------------|--------|---------|
| iPhone 14 | Fingerprint | Wireless | 128GB | <20% |
| iPhone 14 | Face recognition | 3G | 256GB | 20-39% |
| iPhone 14 | Text password | 4G LTE | 512GB | 40-59% |
| iPhone 14 | Fingerprint | 5G | 1TB | 60-79% |
| iPhone 14 | Text password | Wireless | 256GB | 80-100% |
| iPhone 13 | Fingerprint | 3G | 512GB | <20% |
| iPhone 13 | Face recognition | 4G LTE | 1TB | 20-39% |
| iPhone 13 | Text password | 5G | 128GB | 40-59% |
| iPhone 13 | Face recognition | Wireless | 256GB | 60-79% |
| iPhone 13 | Text password | 3G | 1TB | 80-100% |
| Galaxy Z | Fingerprint | 4G LTE | 256GB | <20% |
| Galaxy Z | Face recognition | 5G | 512GB | 20-39% |
| Galaxy Z | Text password | Wireless | 1TB | 40-59% |
| Galaxy Z | Face recognition | 3G | 128GB | 60-79% |
| Galaxy Z | Fingerprint | 4G LTE | 128GB | 80-100% |
| Huawei Mate | Face recognition | 5G | 256GB | <20% |
| Huawei Mate | Text password | Wireless | 512GB | 20-39% |
| Huawei Mate | Fingerprint | 3G | 1TB | 40-59% |
| Huawei Mate | Text password | 4G LTE | 128GB | 60-79% |
| Huawei Mate | Face recognition | 5G | 512GB | 80-100% |
| Google Pixel 7 | Text password | Wireless | 1TB | <20% |
| Google Pixel 7 | Fingerprint | 3G | 128GB | 20-39% |
| Google Pixel 7 | Face recognition | 4G LTE | 256GB | 40-59% |
| Google Pixel 7 | * | 5G | 512GB | 60-79% |
| Google Pixel 7 | * | * | * | 80-100% |

Description of test cases:

The set of 25 test cases generated by the CAGen tool provide comprehensive validation of the mobile application functionality through rigorous combinatorial testing. By leveraging pairwise testing techniques, these tests systematically cover all possible discrete pairs between the input parameters of phone type, authentication mode, network connectivity, memory capacity, and battery level. Each test case represents a unique combination of values across the parameters to

validate both core functions and interactions. Varying these conditions between tests simulates diverse real-world usage scenarios and exposes any flaws or gaps in the application logic.

The key advantage of combinatorial pairwise testing is the ability to provide thorough coverage and maximize defect detection with a minimal and optimized set of test cases. Research has shown pairwise testing to be highly effective at catching interaction faults that are missed by single parameter value testing. By automating the generation of these exhaustive pairwise combinations, CAgen's test cases deliver very high coverage strength and confidence in a compact test suite. Executing these tests validates that all major flows and functions in the mobile application work correctly under different conditions as desired for a high-quality release.

Research and Identifying the Generative AI tool to Use

Researched AI tools for this assignment:

1. ChatGPT
2. BardAI
3. ClaudeAI

Shortlisted tool for this assignment – ***ChatGPT***

Reason:

ChatGPT excels over Claude AI and Bard AI in experimental design and test case generation due to its versatile language understanding, contextual comprehension, continuous knowledge updates, integrations, extensive knowledge base, and interactive nature. Trained on diverse text sources, ChatGPT grasps technical nuances effectively. Its ability to grasp context ensures tailored suggestions. Regular updates maintain relevance in evolving methodologies. Integration capabilities streamline usage in existing workflows. With access to vast information, ChatGPT offers rich insights. Interactivity fosters refinement of requirements. While Claude AI and Bard AI have strengths, ChatGPT's broad coverage, context awareness, updates, integrations, knowledge depth, and interactivity collectively enhance its utility in experimental design and test case generation.

Prompt given to ChatGPT:

Assume that your team has created a new mobile application which requires interaction with a customer to collect user input. The application will require a customer to enter five (5) different types of user inputs with two to five different options. Below is my ACTS configuration file:

[System]

Name: assignment3

[Parameter]

Type of Phone (enum): iPhone 14,iPhone 13, Galaxy Z, Huawei Mate, Google Pixel 7
Authentication (enum): Fingerprint, Face recognition, Text password
Connectivity (enum): Wireless,3G,4G LTE,5G
Memory (enum): 128GB,256GB,512GB,1TB
Battery (enum): <20%,20-39%,40-59%,60-79%,80-100%

Develop sets of test cases using pairwise combination technique (design of experiments)



You

Assume that your team has created a new mobile application which requires interaction with a customer to collect user input. The application will require a customer to enter five (5) different types of user inputs with two to five different options. Below is my ACTS configuration file:

[System]

Name: assignment3

[Parameter]

Type of Phone(enum): iPhone 14,iPhone 13,Galaxy Z,Huawei Mate,Google Pixel 7

Authentication(enum): Fingerprint,Face recognition,Text password

Connectivity(enum): Wireless,3G,4G LTE,5G

Memory(enum): 128GB,256GB,512GB,1TB

Battery(enum): <20%,20-39%,40-59%,60-79%,80-100%

[Constraint]

Develop sets of test cases using pairwise combination technique (design of experiments)



Testcases generated by Generative AI (ChatGPT):

Below is the complete response from ChatGPT to the above prompt:

Using the pairwise combination technique, here are sets of test cases for the given mobile application:

1. Test Cases for Type of Phone and Authentication:

- iPhone 14, Fingerprint
- iPhone 13, Face recognition
- Galaxy Z, Text password
- Huawei Mate, Fingerprint
- Google Pixel 7, Face recognition

2. Test Cases for Type of Phone and Connectivity:

- iPhone 14, Wireless
- iPhone 13, 3G
- Galaxy Z, 4G LTE
- Huawei Mate, 5G

- Google Pixel 7, Wireless

3. Test Cases for Type of Phone and Memory:

- iPhone 14, 128GB
- iPhone 13, 256GB
- Galaxy Z, 512GB
- Huawei Mate, 1TB
- Google Pixel 7, 128GB

4. Test Cases for Type of Phone and Battery:

- iPhone 14, <20%
- iPhone 13, 20-39%
- Galaxy Z, 40-59%
- Huawei Mate, 60-79%
- Google Pixel 7, 80-100%

5. Test Cases for Authentication and Connectivity:

- Fingerprint, Wireless
- Face recognition, 3G
- Text password, 4G LTE

6. Test Cases for Authentication and Memory:

- Fingerprint, 128GB
- Face recognition, 256GB
- Text password, 512GB

7. Test Cases for Authentication and Battery:

- Fingerprint, <20%
- Face recognition, 20-39%
- Text password, 40-59%

8. Test Cases for Connectivity and Memory:

- Wireless, 128GB
- 3G, 256GB
- 4G LTE, 512GB
- 5G, 1TB

9. Test Cases for Connectivity and Battery:

- Wireless, <20%
- 3G, 20-39%
- 4G LTE, 40-59%
- 5G, 60-79%

10. Test Cases for Memory and Battery:

- 128GB, <20%
- 256GB, 20-39%
- 512GB, 40-59%
- 1TB, 60-79%

These test cases cover pairwise combinations of the different types of user inputs, ensuring comprehensive testing while minimizing the number of test cases required.

Screenshots:



ChatGPT

Using the pairwise combination technique, here are sets of test cases for the given mobile application:

1. Test Cases for Type of Phone and Authentication:
 - iPhone 14, Fingerprint
 - iPhone 13, Face recognition
 - Galaxy Z, Text password
 - Huawei Mate, Fingerprint
 - Google Pixel 7, Face recognition
2. Test Cases for Type of Phone and Connectivity:
 - iPhone 14, Wireless
 - iPhone 13, 3G
 - Galaxy Z, 4G LTE
 - Huawei Mate, 5G
 - Google Pixel 7, Wireless
3. Test Cases for Type of Phone and Memory:
 - iPhone 14, 128GB
 - iPhone 13, 256GB
 - Galaxy Z, 512GB
 - Huawei Mate, 1TB
 - Google Pixel 7, 128GB
4. Test Cases for Type of Phone and Battery:
 - iPhone 14, <20%
 - iPhone 13, 20-39%
 - Galaxy Z, 40-59%



After this response I asked ChatGPT to generate the complete set of 25 test cases for this experiment. Below is the screenshot of the tabular form of 25 test cases given as output:



Certainly! Here are all 25 test cases generated using pairwise combination for the given parameters:

| Test Case | Type of Phone | Authentication | Connectivity | Memory | Battery |
|-----------|----------------|------------------|--------------|---------|---------|
| 1 | iPhone 14 | Fingerprint | Wireless | 128GB | <20% |
| 2 | iPhone 13 | Face recognition | 3G | 256GB | 20-39% |
| 3 | Galaxy Z | Text password | 4G LTE | 512GB | 40-59% |
| 4 | Huawei Mate | Fingerprint | 5G | 1TB | 60-79% |
| 5 | Google Pixel 7 | Face recognition | Wireless | 128GB | 80-100% |
| 6 | iPhone 13 | Text password | 4G LTE | 1TB | 20-39% |
| 7 | Galaxy Z | Fingerprint | 3G | 256GB | <20% |
| 8 | Huawei Mate | Face recognition | 5G | 512GB | 60-79% |
| 9 | Google Pixel 7 | Text password | Wireless | 512GB | 40-59% |
| 10 | iPhone 14 | Face recognition | 5G | 256GB | 20-39% |
| 11 | Galaxy Z | Fingerprint | 4G LTE | 1TB | 80-100% |
| 12 | Huawei Mate | Text password | Wireless | 128GB | 60-79% |
| 13 | Google Pixel 7 | Fingerprint | 3G | 512GB | 20-39% |
| 14 | iPhone 14 | Text password | 3G | 512GB | 40-59% |
| 15 | Galaxy Z | Face recognition | 5G | 128GB | 60-79% |
| 16 | Huawei Mate | Fingerprint | Wireless | 256GB | 80-100% |
| 17 | Google Pixel 7 | Text password | 4G LTE | 256GB | <20% |
| 18 | iPhone 13 | Fingerprint | Wireless | 512GB | 60-79% |
| 19 | Galaxy Z | Text password | 3G | 80-100% | 20-39% |
| 20 | Huawei Mate | Face recognition | 4G LTE | 512GB | 40-59% |
| 21 | Google Pixel 7 | Fingerprint | 5G | 1TB | <20% |
| 22 | iPhone 13 | Text password | Wireless | 256GB | 40-59% |
| 23 | Galaxy Z | Face recognition | 3G | 80-100% | 60-79% |
| 24 | Huawei Mate | Fingerprint | 4G LTE | 256GB | 20-39% |
| 25 | Google Pixel 7 | Text password | 5G | 128GB | 40-59% |

Description of test cases:

1. Test Case 6:

- Type of Phone: iPhone 13
- Authentication: Text password
- Connectivity: 4G LTE

- Memory: 1TB
 - Battery: 20-39%
 - Description: This test case represents a scenario where the user is using an iPhone 13 with text password authentication, connected to a 4G LTE network, with a large memory capacity of 1TB, and the battery level is between 20% and 39%.
2. **Test Case 12:**
- Type of Phone: Huawei Mate
 - Authentication: Text password
 - Connectivity: Wireless
 - Memory: 128GB
 - Battery: 60-79%
 - Description: In this test case, the user is utilizing a Huawei Mate device, employing text password authentication, connected wirelessly, equipped with 128GB of memory, and the battery level falls between 60% and 79%.
3. **Test Case 17:**
- Type of Phone: Google Pixel 7
 - Authentication: Fingerprint
 - Connectivity: 3G
 - Memory: 512GB
 - Battery: 20-39%
 - Description: This scenario involves a user operating a Google Pixel 7 smartphone, utilizing fingerprint authentication, connected to a 3G network, having a large memory capacity of 512GB, and the battery level is between 20% and 39%.

Assessment of tests developed by CAgen (DOE tool):

Validity Assessment:

The 25 test cases generated by CAgen provide highly valid combinatorial coverage for the mobile app parameters. By covering all pairwise combinations, they create a diverse and robust set of test conditions focused on interactions.

1. Thorough Parameter Coverage: All the defined parameters - phone type, authentication, connectivity, memory, and battery level are comprehensively covered in the test cases.
2. Pairwise Interaction Coverage: The test cases effectively cover discrete pairs between all parameters through different value combinations to fully validate interactions.
3. Real-World Usage Simulation: Varying values in diverse ways simulates real user scenarios for thorough validation.

Alignment with DOE Guidelines:

The pairwise test generation approach is well aligned with fundamental DOE techniques and objectives.

1. Combination Coverage: Guidelines emphasize exploring combinations of factor levels, which pairwise testing systematically achieves.
2. Interaction Focus: The tests focus on uncovering interaction faults, a key goal of DOE.
3. Optimized Test Suite: Pairwise testing minimizes tests by optimizing the combination coverage as per DOE principles.
4. Repeatable Tests: The documented test cases allow repeatability and reproducibility, essential for DOE.

In summary, the CAgen tests exhibit strong validity in terms of the coverage provided and align well with core DOE criteria for interaction testing using optimized test suites. This assessment highlights the effectiveness of the tool and the technique.

Assessment of tests developed by ChatGPT (Generative AI):

Validity Assessment:

The 25 test cases from ChatGPT provide broad coverage but lack the rigorous combinatorial approach seen in CAgen.

1. Parameter Coverage: All parameters are included in the test cases.
2. Value Coverage: Most values are covered but there are some gaps like 80-100% battery level missing, and is instead added to the battery column, generating wrong test cases.
3. Interaction Coverage: Test cases focus more on covering individual values rather than value pairs. Lacks structured pairwise coverage.
4. Test Quality: Some tests have same values repeated while some valid combinations are missed indicating suboptimal generation.

Alignment with DOE Guidelines:

The ChatGPT tests have some alignment but lack the optimization and combination focus core to DOE.

1. Coverage Breadth: Tests do explore combinations of factor levels broadly though not exhaustively.
2. Interaction Testing: Limited focus on uncovering interaction faults.
3. Test Suite Optimization: Test suite seems arbitrarily generated rather than optimized for efficiency.
4. Repeatability: Documented tests allow repeatability which aligns with DOE.

In summary, while the ChatGPT tests provide reasonable validity and alignment, they lack the structured pairwise coverage and optimization critical for rigorous DOE testing. The CAgen test cases are more aligned with DOE objectives. But ChatGPT shows potential for further enhancement.

Assessment of CAgen (DOE tool)

Features and Functionalities:

1. CAgen provides an intuitive web-based interface for combinatorial test design without any complex setup.
2. It supports various test generation algorithms like pairwise, 3-wise coverage to choose from.
3. The tool allows uploading ACTS file, configuring parameters, generating tests, and exporting results seamlessly.
4. It generates optimal minimized test suites through built-in optimization.
5. Features like test case previews, integrated reporting make CAgen powerful.

Scope Covered:

1. CAgen covers combinatorial testing approaches like pairwise, n-wise testing very effectively.
2. It is designed specifically for parameter-based combinatorial testing of software systems.
3. Within combinatorial testing, CAgen provides comprehensive capabilities.

Performance:

1. It generates optimized test suites in a very fast and efficient manner.
2. Test case generation scales smoothly even for large configurations.
3. The results deliver high quality tests with minimal length.

Ease of Use:

1. The tool provides an intuitive and user-friendly web interface requiring no training.
2. Creating tests by uploading ACTS files is very straightforward.
3. Exporting test cases to various formats for execution is seamless.
4. CAgen makes applying advanced combinatorial testing easy.

In summary, CAgen delivers a full-featured, high performance, user-friendly combinatorial testing solution that performed very well for this assignment. Its capabilities and usability make CAgen a great choice for combinatorial testing needs.

Assessment of ChatGPT (Generative AI tool)

ChatGPT Assessment for Design of experiments and generating test cases:

1. ChatGPT was able to generate reasonable test cases when provided the parameters and guidance through prompts.
2. However, the test cases lacked rigorous combinatorial coverage and optimization seen with specialized tools like CAgent.
3. More complex prompt engineering is needed to get ChatGPT to generate structured pairwise or n-wise test suites.
4. There is potential to enhance ChatGPT's capabilities further with more contextual training in combinatorial testing.
5. Overall, in its current state, it is not a replacement for commercial DOE tools but shows promise as an assistive tool.

Significance of Generative AI for DOE:

1. Automated test case generation from parameters and configurations can save significant effort.
2. Generative AI like ChatGPT can provide reasonably good test cases faster compared to manual creation.
3. However, achieving true combinatorial coverage and optimization currently needs specialized tools.
4. With further training on testing concepts, Generative AI could become more proficient at DOE techniques.
5. It has potential to make combinatorial testing more accessible to testers without expertise.
6. Prompt engineering will play a key role in leveraging Generative AI optimally for DOE.

In summary, while ChatGPT shows potential, currently combinatorial DOE requires specialized tools. But Generative AI presents possibilities to assist and enhance test design productivity, if guided properly through robust prompt engineering. More contextual training could unlock its true potential for DOE.