

テストプロセスにおける「ChatGPT」の活用の提案

堀川 尚人
株式会社HBA
horikawaxn@hba.co.jp

要旨

本稿では、膨大かつ複雑化するシステムテストにおける課題を解決するために人工知能技術の ChatGPT を活用する方法を提案する。

ChatGPT を活用して、テスト担当者の負荷を軽減し、高品質なテストを実現することが可能か、テストプロセスの一部をサポートさせることが可能かを検証し、その有効性や可能性について議論したい。

1. はじめに

システムテストは、ソフトウェア開発においてシステム全体の品質を確保するための重要なフェーズである。システムテストを効果的に実施することは、高いソフトウェア品質を達成するために必要不可欠である。しかし、テストプロセスはしばしば時間がかかる。

筆者は車載開発に関するテスト業務に関わっているため自動車を例に挙げると、自動車に搭載する機能が年々増加かつ複雑化しており、それによってテストケースも増加かつ複雑化している。

今後の車載開発では「機能や複雑度が増加したにもかかわらず、エンドユーザーは自動車の機能に対して従来と同等またはそれ以上の品質」[1]を求めるため、今後さらに増加かつ複雑化するテストケースを、従来と同様の期間で作成できるようにすることが今後の課題となる。これは車載開発に関わらず他の多くのシステム開発でも同様であると考えられる。

また、テスト品質はテスト担当者の経験にも依存するため、高品質なテストを実施するには経験豊富な人材が必要となる。しかしソフトウェアテストの経験や知識に富んだ人材はなかなかおらず、人材育成にも時間と予算が必要となる。

そこで、これらの問題を解決するために人工知能に注目した。

近年の人工知能の進歩により、自然言語処理技術を

活用した ChatGPT が注目されている。ChatGPT は、人工知能によって学習された言語モデルであり、自然言語処理技術を用いて文章の生成や応答を行うことができる。

テストプロセスにおいて求められる能力の一つに、日本語で書かれた要求を分析することが挙げられる。そこで、ChatGPT を用いてテストプロセスの一部をサポートできれば、テスト担当者の負荷軽減や高品質なテストの実現が期待できると考えられる。ただし、大規模開発に ChatGPT を活用した具体的な事例はなく、テストプロセスに ChatGPT を活用した事例も見つけることができなかった。

そのため、本稿では小規模な事例を用いてテストプロセスにおける ChatGPT の活用可能性を検証し、テストプロセスをサポートする方法とその有効性を提案する。

2. 検証方法

本稿で検証するテストタイプとテストレベルは、テストレベルがシステムテスト、テストタイプは機能テストを想定した。

検証方法として、数種類の要求サンプルを用意し、それらに対して ChatGPT に質問を投げ、その出力結果から有用性を検証した。ChatGPT のモデルは GPT-4 を使用した。なお、本検証では ChatGPT の回答の正確性に焦点を当てるのではなく、ChatGPT が実現できる機能を検証することに重点を置いている。

検証方法の例は図 1 である。要求サンプルの一つに、マイヤーズの三角形問題[2]を使用したため、それを用いたケースである。マイヤーズの三角形問題とは、ユーザーが入力した 3 つの整数を三角形の辺の長さとして、不等辺三角形、二等辺三角形、正三角形のうちどれに該当するかを判定し、結果をメッセージで出力するプログラムに対して、テストケースを導出する問題である。

以下、このプログラムをテストするのに、必要なテスト観点を書いてください。
なお、具体的な値は書かないでください。

このプログラムでは、ユーザーが3個の整数を入力します。
この3個の値は、それぞれ三角形の3辺の長さを表すものとします。
プログラムは、三角形が、不等辺三角形、二等辺三角形、正三角形のうちのどれであるかを定めるメッセージを出力します。

以下に、このプログラムをテストする際に考慮すべきテスト観点を示します。

1. **正常な入力値**：三角形を形成するための正常な3つの整数値を入力する。これは、プログラムが正しく三角形のタイプを判断できるかをテストします。
2. **不正な入力値**：三角形を形成できない3つの整数値を入力する。これは、プログラムが無効な入力を適切に扱えるかをテストします。
3. **正三角形のテスト**：3辺すべてが等しい値の入力をテストします。これは、プログラムが正三角形を正確に判断できるかを確認します。
4. **二等辺三角形のテスト**：2辺が等しく、3番目の辺が異なる値の入力をテストします。これは、プログラムが二等辺三角形を正確に判断できるかを確認します。
5. **不等辺三角形のテスト**：3辺すべてが異なる値の入力をテストします。これは、プログラムが不等辺三角形を正確に判断できるかを確認します。

図 1

このように要求サンプルに対して、ChatGPT にテストプロセスの様々な質問を繰り返し、ChatGPT の回答からその活用の可能性を検討した。

3. 提案

本章では、各テストプロセスにおいて、ソフトウェアテストにおける ChatGPT の活用可能性を示していく。

3.1. テスト分析

ChatGPT は、言語処理技術に優れているため、テスト分析において要件仕様書やドキュメントを解析し、解析した結果をアウトプットすることができると考えられる。例えば、要求仕様書を解析することで、テスト観点の抽出が可能だ。

実例としては、図 1 のように要求を記載してテスト観点の抽出する旨を求めると、ChatGPT はテスト観点の代表的な例を出力した。あくまで代表的な例であるため、出力した内容をそのまま使用することはできないが、ヒントとしては利用できるであろう。

また、要求仕様書の曖昧な点、矛盾点、考慮点の抽出も可能だ。要求仕様書には多数の要求が含まれており、その要求の不備や内容の矛盾が含まれる場合がある。そのケースを抽出できるかどうかとも検証した。

実例としては図 2 の通り、マイヤーズの三角形問題に対して、曖昧な点、矛盾点、仕様として漏れている点の質問をしたところ、入力値の範囲、ユーザーが整数をどのような形式で入力するか、ユーザーが無効な入力をした場合のエラー処理が仕様に記載されていないことを挙げた。

このように ChatGPT を使うことで、仕様の漏れやあいまいな点といった要求仕様書の不備を自動で抽出することができるため、要求仕様書の品質向上にもつなげられるであろう。

以下はプログラムに関する仕様です。これらに対して、曖昧な点、矛盾点、仕様として漏れている点を挙げてください。

このプログラムでは、ユーザーが3個の整数を入力します。
この3個の値は、それぞれ三角形の3辺の長さを表すものとします。
プログラムは、三角形が、不等辺三角形、二等辺三角形、正三角形のうちのどれであるかを定めるメッセージを出力します。

この仕様に関して、以下の曖昧な点、矛盾点、仕様として漏れている点が挙げられます。

1. 入力値の範囲: 仕様では整数が入力されることが明記されていますが、入力される整数の範囲について言及されていません。例えば、負の整数や非常に大きな整数が入力された場合、プログラムはどのように対処すべきか明確ではありません。
2. 三角形の成立条件: 仕様には、三角形の成立条件についての言及がありません。三角形の成立条件とは、任意の2辺の和が残りの1辺よりも大きいという条件です。この条件が満たされない場合、プログラムはどのようなメッセージを出力すべきか明確ではありません。
3. 入力方法: ユーザーが整数をどのような形式で入力するか、例えばコマンド区切り、空白区切り、またはそれぞれ別々の行で入力するかなど、明確にされていません。入力方法の明確化が必要です。
4. エラー処理: 仕様には、ユーザーが無効な入力をした場合のエラー処理についての記載がありません。例えば、整数以外の値が入力された場合や、入力不足している場合にどのようなエラーメッセージを表示するかなど、明確化が必要です。

図 2

3.2. テスト設計

テスト設計においても、テスト分析と同様に、要求仕様書やテスト分析結果を ChatGPT にインプットとして使用することで、テストプロセスのサポートが可能と考えている。

例えば、テスト分析で挙げたテスト観点に対して、ChatGPT に網羅の基準とその網羅理由を挙げさせることができる。実例の一部が図 3 となる。

図 3 では、マイヤーズの三角形問題に対して事前に出力していた、「すべての辺の長さが正の整数である場合、正しい種類の三角形が出力されることを確認する」というテスト観点に対して、網羅の基準とその網羅理由を挙げる指示を出したところ、網羅の基準として「不等辺三角形、二等辺三角形、正三角形の 3 種類の三角形が出力されること」を挙げ、網羅理由は「三角形の種類を正確に判別できることを確認するために、各種類について最低 1 つの例をテストする。」という内容を挙げた。

1. すべての辺の長さが正の整数である場合、正しい種類の三角形が出力されることを確認する。
 - 網羅の基準:
 - a. 不等辺三角形: 辺の長さがそれぞれ異なる (例: 3, 4, 5)
 - b. 二等辺三角形: 辺の長さが2つだけ等しい (例: 3, 3, 5)
 - c. 正三角形: すべての辺の長さが等しい (例: 3, 3, 3)
 - 理由: 三角形の種類を正確に判別できることを確認するために、各種類について最低1つの例をテストする。

図 3

さらに、要求から適切なテスト設計手法を出力できるかどうかを検証した。

この検証に使用した要求サンプルは、エアコンの仕様を簡素化したもので、「温度を 1℃ずつ変えることができ、冷房運転時の設定範囲は 20 度～31 度、暖房運転時の設定範囲は 16 度～25 度」という内容だ。

この仕様に対してテスト設計手法を考えるように指示したところ、「境界値分析、同値分割、状態遷移テスト」の 3 つのテスト設計手法を当てはめられることを挙げた。

このように、ChatGPT が出力した網羅の基準やテスト設計手法などのテスト設計に必要な情報をヒントとして活用することで、自分で一から作成する場合よりテスト担当者の負担や時間をかけずにテスト設計を行うことができるであろう。

3.3. テスト実装

テスト実装フェーズでは、ChatGPT をテストケースの作成に活用できると考えている。ChatGPT は、テスト設計のアウトプットである「何をどのようにテストするか」という内容に基づいて、テストケースを生成することができる。

実例として図 3 の内容をもとに、ChatGPT でテストケースを出力した内容が図 4 である。図 3 で出力した三角形の網羅基準をもとに、入力値と出力値をテストケースとして ChatGPT が出力している。

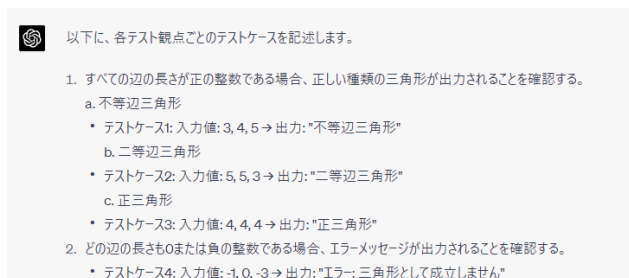


図 4

さらに図 4 で出力したテストケースに対してバリエーションを増やすよう指示をすると、入力値が異なるテストケースを ChatGPT は出力した。

以上から ChatGPT はテスト観点をインプットすることで、図 4 のようにその観点に基づいたテストケースを出力することができる。テストケースが足りない場合は、類似のテストケースを出力する指示を ChatGPT にすることで、その指示に従ったテストケースの出力もできる。

このことから既存のテスト仕様書をインプットにテストケースの拡張も対応でき、テスト設計工程で作成したテスト観点があればテストケースの生成も可能であろう。

また、テストケースの入力内容(前提条件、テスト手順、期待値等)を指定することで、テスト担当者が期待したテストケースを生成することも可能だ。図 4 のテストケースの入力内容を指定したケースが図 5 である。

テスト担当者が求めているテストケースのフォーマットがある場合は、それらを指定することで ChatGPT の出力したデータを形成する手間も省くことができるであろう。

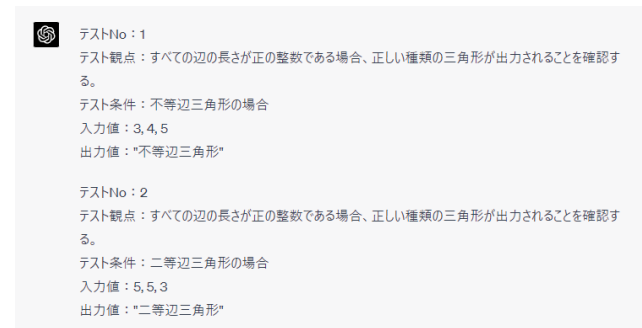


図 5

以上のように、ChatGPT を用いることで、テストプロセスの多くの部分をサポートし、テスト担当者の負荷を軽減することができる可能性が見えてきた。

ただし、ChatGPT の精度には限界があるため、テスト担当者は自らの知識や経験を持って、ChatGPT の回答を適切に評価する必要がある。もしテストプロセスに取り込むのであれば、全プロセスを ChatGPT に置き換えるのではなく、各プロセスの一部を ChatGPT に置き換えるところから始めるべきであろう。

4. 課題

本稿の検証内容に限った話ではないが、ChatGPT を活用する際にはいくつかの課題が存在する。

前述の内容と重複するが ChatGPT の精度には限界があるため、テスト担当者が ChatGPT の回答を適切に評価する必要がある。ChatGPT は質問に対して人間らしい回答をするが、情報の正確性や信頼性は保証されていない。人間が作成した成果物と同様に、ChatGPT の成果物に対してもレビューが必要となるであろう。

さらに、情報漏洩等のセキュリティ上の問題も懸念される。ChatGPT の FAQ に「API 経由で送られたお客様のデータを弊社のモデル改良のために使用することはありませぬ。」[3]とあるため、API 経由であればこの問題は解決できそうだが、現状は機密情報や個人情報ではないデータを用いるべきであろう。

5. おわりに

本稿では、ChatGPT を使用してソフトウェアテストのプロセスをサポートする方法を提案し、その有効性を検証した。

結果として、ChatGPT を使用することで、テストケースの作成やテスト観点の抽出、顧客要求の曖昧な点や矛盾点の特定などの作業のサポートできる可能性が見えてきた。

ただし、ChatGPT はリリースされて間もないサービスである。活用やセキュリティなど解決すべき課題があるが、ChatGPT の可能性はまだまだ秘めていると考えているため、今後の進歩に期待したい。

参考文献

- [1] ISTQB テスト技術者資格制度
Foundation Level Specialist シラバス 自動車ソフトウェアテスト担当者 日本語版 Version 2018.J03
https://jstqb.jp/dl/JSTQB-SyllabusFoundation_Version2018V31.J03.pdf
- [2] Glenford J. Myers, ソフトウェア・テストの技法 第 2 版, 近代科学社, 2006
- [3] ChatGPT APIFAQ
<https://help.openai.com/en/articles/7039783-chatgpt-api-faq>