

単体テスト生成における ベテラン思考の大規模言語モデルへの適用

キャノンイメージングシステムズ株式会社

小林 智也

単体テスト作成における課題

ソフトウェア保守開発における単体テストの課題

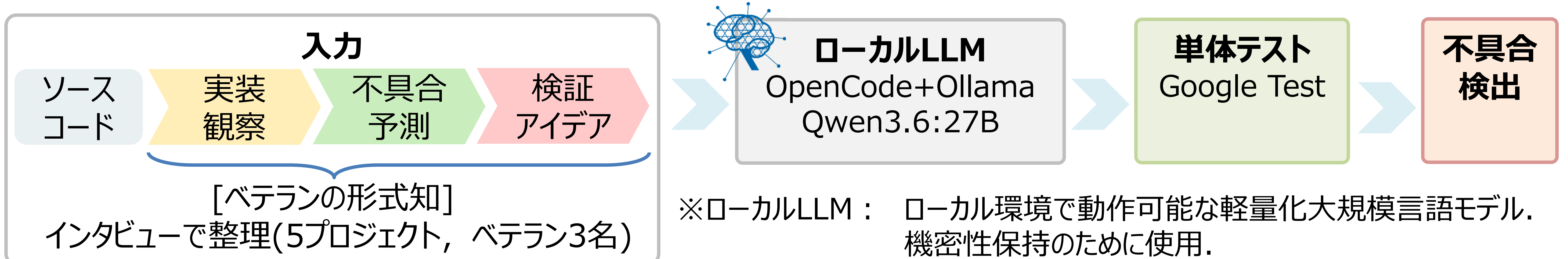
- 十分な設計ドキュメントが得られないケースでは、単体テストの作成コストが高い。
- 高品質な単体テスト作成は、ベテラン技術者に依存。ベテラン技術者の継続的な確保は困難。

手法・ツールの適用による解決

ベテラン思考を形式化し大規模言語モデルに適用

- 単体テスト作成時の思考を分析し、ベテランの形式知として整理。
- 整理した形式知を大規模言語モデルに適用し、生成した単体テストの不具合検出能力を分析。

システム概要



検証アプローチと結果

[実案件の不具合を用いて評価]

- C1** 大量の似た処理に紛れた、setter設定漏れ。一般的な不具合。
- C2** 画像の表示不正。専門知識が必要。

[LLMに与える形式知の組み合わせと不具合検出結果]

					C1	C2	知見
ソースコード				ローカルLLM	×	×	1
ソースコード	+	A	+	ローカルLLM	✓	✓	
ソースコード			+	ローカルLLM	✓	×	2
ソースコード	+	A		ローカルLLM	✓	×	3
					+ 検証アイデア		
					A: 実装観察 B: 不具合予測		

得られた知見

- LLMだけでは検出できない不具合を **A** **B** の両方を与えることで検出。
- 一般的な設定漏れは、**A** **B** の片方を与えるだけで検出。形式知の一部でも不具合を具体化。
- 専門知識が必要な不具合は、**A** **B** の片方だけでは検出が出来ない。実装の手掛かり×専門知識が必要。

結論と今後

結論

ベテラン思考の形式知化により、不具合検出能力のある単体テスト生成を支援できる可能性を示した。

今後

- 一般化を目的とした実験不具合の拡充。実験回数を増やした統計的検証。
- 不具合以外のベテラン形式知も与える。単体テスト生成にどのような影響を与えるか、検証。