

コードクローン検出ツールCCFinderNexGenの開発

開発者: 神谷年洋

- あなたのソースコード、大丈夫ですか？
 - ☑「動いているコードに触るな」方針で何年も保守されてきた
 - ☑最近、修正のためのコストが上昇している
 - ☑外注したコードの中身がチェックできていない
- 一つでもあてはまれば、コードクローン分析が必要です！

コードクローン問題

コードクローンとは、ソースコードの重複部分。

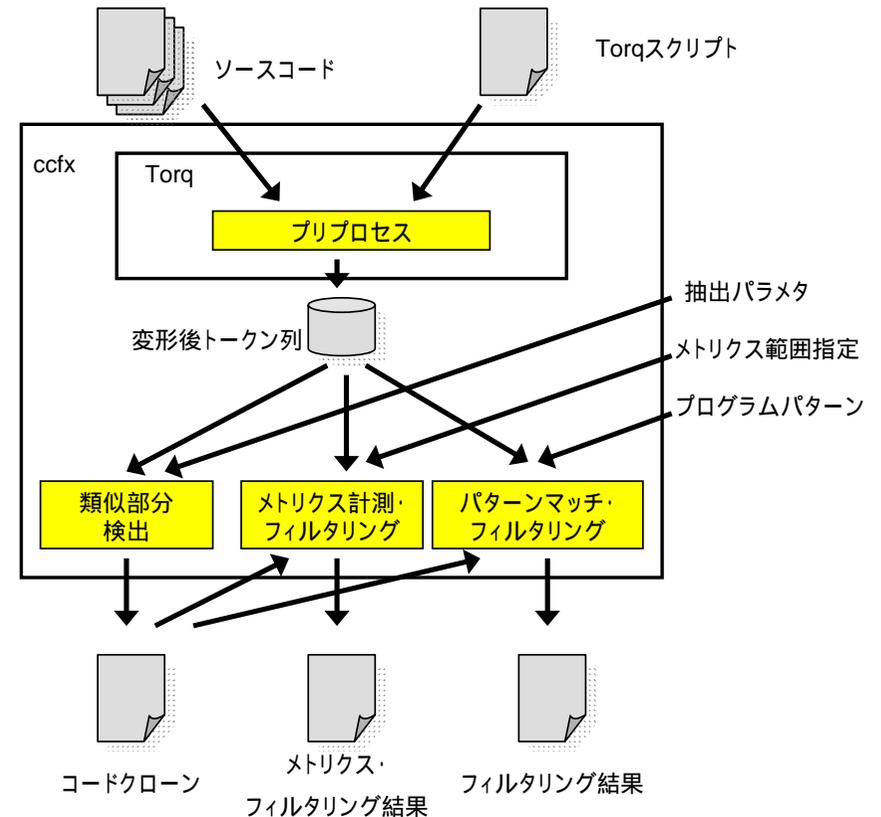
- 典型的には、ソースコードのコピー & ペーストによって生じる
- 大規模ソフトウェアの50%がコードクローンだったという事例も)

ソフトウェア保守の大問題

- 「重複したロジックは修正が困難である」Kent Beck(リファクタリングの第一人者)
- GPLコードからのコピー & ペーストが発覚すれば、知財問題に

コードクローン対策の強力な助っ人 CCFinderX

- CCFinderXは、コードクローンを検出・分析するためのツールです
- 強力な新検出エンジン
 - 大規模な(数百万行～)ソースコードに対応
 - 変数名の付け替えなどが行われたコードも検出
- 新機能: プリプロセスのカスタマイズ
 - プログラミング言語の方言に対応
 - 新しいプログラミング言語に対応
- 他のツールと連携
 - CSV形式のファイルを出力
 - Eclipseへの組み込み(予定)



- 分析機能
 - 散布図による直感的な把握とメトリクスによる定量的な分析

散布図:
コードクローンの分布を直感的に把握

ソーステキスト:
コードクローンになっているソースコード断片を参照

The screenshot displays the GemX application interface. On the left, a table lists file metrics. In the center, a scatter plot visualizes the distribution of code clones. On the right, source code snippets are shown with red highlights indicating cloned code.

File ID	Path	Length	# Clones	NBR	RSA00	RS1%
1443	ndWmnd.c	16259	0	0	0.00	0.00
1444	ndWmktables.c	496	0	0	0.00	0.00
1445	ndWmktables.c	2531	6	4	22.48	0.00
1446	ndWmktables.c	2534	1	1	5.07	0.00
1447	ndWmktables.c	6542	15	4	96.93	0.00
1448	ndWmktables.c	8362	16	5	32.04	0.00
1449	ndWmktables.c	10698	21	4	76.42	1.98
1450	ndWmktables.c	339	0	0	0.00	0.00
1451	ndWmktables.c	11630	20	4	70.26	0.00
1452	ndWmktables.c	681	4	2	74.16	0.00
1453	ndWmktables.c	544	0	0	0.00	0.00
1454	ndWmktables.c	855	8	2	94.85	0.00
1455	ndWmktables.c	1541	8	2	58.21	13.11
1456	ndWmktables.c	512	0	0	0.00	0.00
1457	ndWmktables.c	413	0	0	0.00	0.00
1728	messageFusionWinx_compath	18	0	0	0.00	0.00
1729	messageFusionWinx_compath	1021	0	0	0.00	0.00
1730	messageFusionWinx_compath	685	1	0	0.00	24.23
1731	messageFusionWinx_compath	6	0	0	0.00	0.00
1732	messageFusionWinx_compath	17	0	0	0.00	0.00
1733	messageFusionWinx_compath	23151	10	1	1.01	7.24
1734	messageFusionWinx_compath	12146	16	0	0.00	12.79
1735	messageFusionWinx_compath	7011	2	0	0.00	7.05
1736	messageFusionWinx_compath	23861	2	1	0.98	0.00
1737	messageFusionWinx_compath	173	16	0	0.00	96.53
1738	messageFusionWinx_compath	418	1	0	0.00	32.06
1739	messageFusionWinx_compath	2041	4	2	13.08	0.00
1740	messageFusionWinx_compath	507	5	3	93.29	0.00
1741	messageFusionWinx_compath	350	2	1	61.43	0.00
1742	messageFusionWinx_compath	339	1	1	40.12	0.00
1743	messageFusionWinx_compath	1062	1	0	0.00	19.59
1744	messageFusionWinx_compath	223	1	1	51.57	0.00
1745	messageFusionWinx_compath	285	1	1	47.72	0.00
1746	messageFusionWinx_compath	183	0	0	0.00	0.00
1747	messageFusionWinx_compath	550	2	2	19.82	0.00
1748	messageFusionWinx_compath	452	2	1	46.24	0.00
1749	messageFusionWinx_compath	23	0	0	0.00	0.00
1750	messageFusionWinx_compath	97	0	0	0.00	0.00
1751	messageFusionWinx_compath	128	0	0	0.00	0.00
1752	messageFusionWinx_compath	1785	26	0	0.00	42.24
1753	messageFusionWinx_compath	2299	0	0	0.00	0.00
1754	messageFusionWinx_compath	1238	2	0	0.00	16.56
1755	messageFusionWinx_compath	1428	0	0	0.00	0.00
1756	messageFusionWinx_compath	4215	0	0	0.00	0.00
1757	messageFusionWinx_compath	5562	4	0	0.00	33.62

ファイルメトリクス:
ソースファイルをコードクローンにより評価。「このファイルの何%がコードクローン？」

クローンセットメトリクス:
コードクローンの特徴を定量的に把握。「このコードクローンはディレクトリ上でどのくらい広がっているの？」