コードクローン検出・変更管理ツール群の開発

本田 紘貴[1] 吉田 則裕[2] 井上 克郎[1]　1 大阪大学　2 名古屋大学

Software Engineering Laboratory

多言語に対応可能なクローン検出ツールCCFinderSW [1]

概要

- 字句解析ベースのクローン検出器CCFinderXは5つの言語に対応
- 対応言語を増やすためには個別に字句解析の実装が必要

　問題：それぞれの字句解析の実装には手間がかかる

➡ 対応したい言語の文法を容易に入力できる仕組みを作成

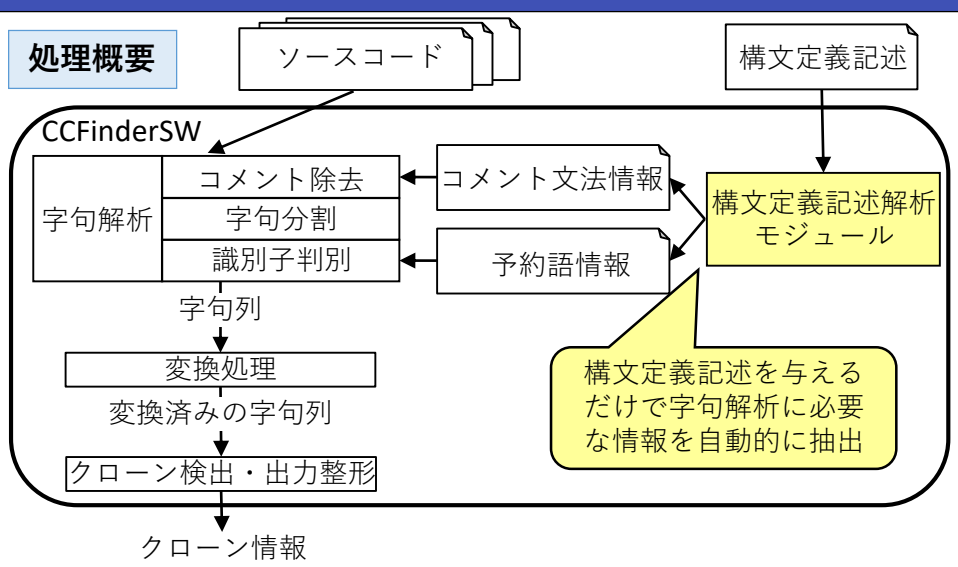
構文定義記述の利用

- ANTLR[1]は構文定義記述から字句解析器や構文解析器を生成
- Grammars-v4[2]はANTLRで利用可能な150以上の構文定義記述を所有

➡ 構文定義記述から字句解析に必要な情報（コメント文法，予約語，文字列リテラル）を自動抽出するモジュールの開発

[1]http://www.antlr.org/　[2]https://github.com/antlr/grammars-v4

処理概要



評価

- 42言語のうち83％の言語ですべての必要な情報を抽出できた
- C++に対する検出において，CCFinderXと同等の検出能力を確認

コードクローン変更管理システムClone Notifier [2]

背景

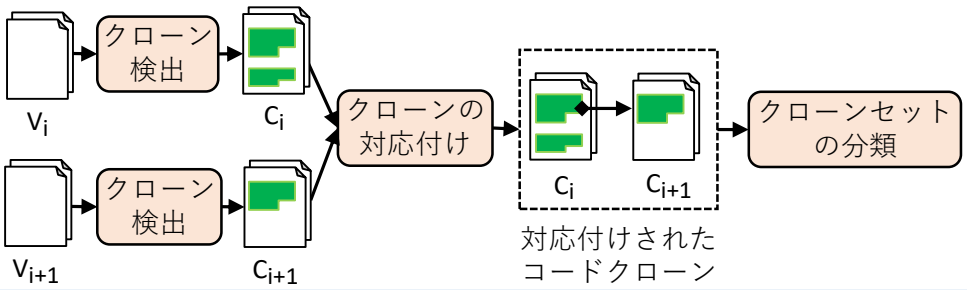
クローンセット：互いにクローンとなっているコード片の集合

クローンセットに対する保守作業

- 同時修正：クローンセット中のコード片を一貫して修正すること
- 集約　　：クローンセット中のコード片を1つにまとめ，サブルーチンで呼び出すこと

➡ 保守作業の効率化のためにコードクローンの変更管理が必要

処理概要



クローンセットの分類

- Stable　：2バージョン間に存在し，変更されなかったもの．
- New　　：新バージョンのみに存在するもの．集約対象の候補．
- Deleted　：旧バージョンのみに存在するもの．集約完了を確認可．
- Changed：2バージョン間に存在し，変更されたもの．同時修正されているか確認可．

➡ 保守対象となる可能性の高いクローンセット情報を提供

評価

企業のソフトウェア開発（保守段階）に約40日間適用

対象プロジェクト

- 開発言語：Java
- 開発規模：約120万行
- 開発人数：6人
- クローンセット数：約850

開発者は，保守対象クローンセット11個を発見できた

コードクローン変更履歴可視化システムCCEvovis [3]

概要

Clone Notifierは，2バージョン間の変更履歴情報のみを提供

　問題：初心者の開発者にとっては使いにくい

　- どの2バージョンを入力すればよいかわからない

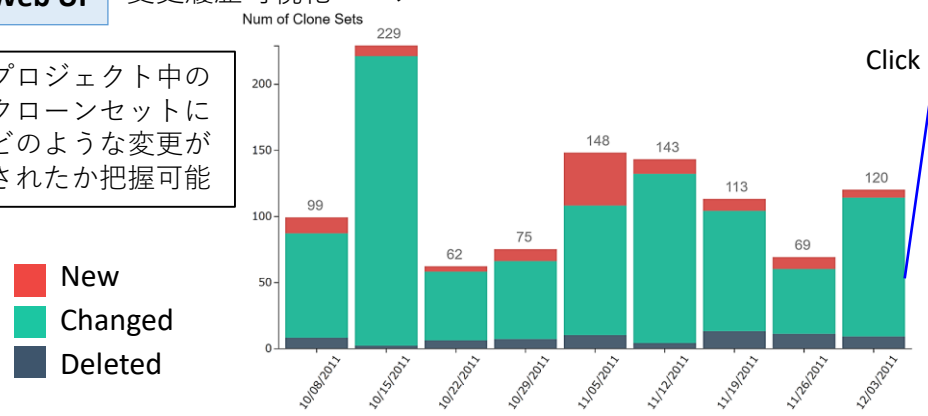
　- 重要な変更履歴を見落とす可能性がある

➡ 複数バージョン間の変更履歴全体を分析し可視化

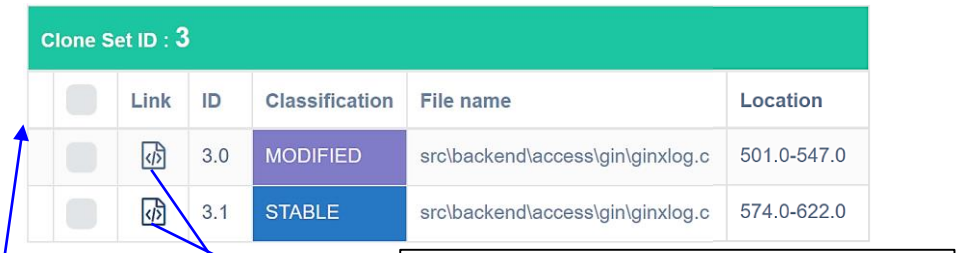
Web UI

変更履歴可視化ページ

プロジェクト中のクローンセットにどのような変更がされたか把握可能



クローンセット一覧ページ



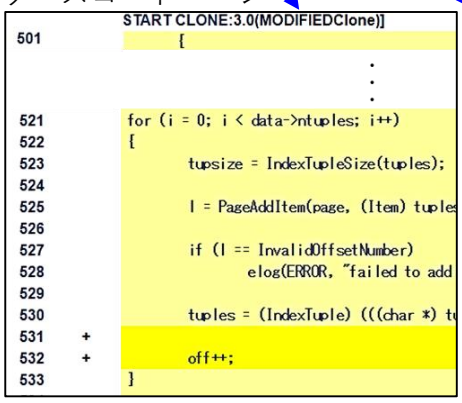
Clone Set ID : 3

Link	ID	Classification	File name	Location
	3.0	MODIFIED	src\backend\access\gin\ginxlog.c	501.0-547.0
	3.1	STABLE	src\backend\access\gin\ginxlog.c	574.0-622.0

クローンセット内に修正されたクローンと修正されていないクローンが存在

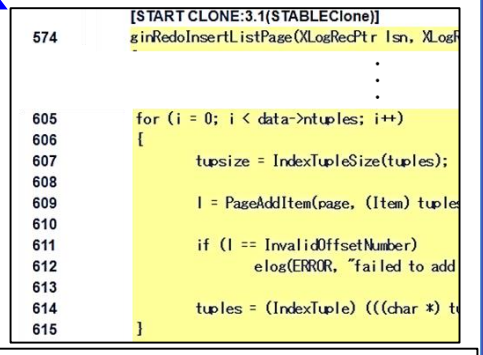
Click

ソースコードページ



START CLONE:3.0(MODIFIEDClone)]

```
501        {
                    :
                    :
521        for (i = 0; i < data->ntuples; i++)
522        {
523            tupsize = IndexTupleSize(tuples);
524
525            l = PageAddItem(page, (Item) tuple
526
527            if (l == InvalidOffsetNumber)
528                elog(ERROR, "failed to add
529
530            tuples = (IndexTuple) (((char *) t
531    +
532    +       off++;
533        }
```



[START CLONE:3.1(STABLEClone)]

```
574        ginRedoInsertListPage(XLogRecPtr lsn, XLogR
                    :
                    :
605        for (i = 0; i < data->ntuples; i++)
606        {
607            tupsize = IndexTupleSize(tuples);
608
609            l = PageAddItem(page, (Item) tuple
610
611            if (l == InvalidOffsetNumber)
612                elog(ERROR, "failed to add
613
614            tuples = (IndexTuple) (((char *) t
615        }
```

同時修正されていないコード片発見

URL　CCFinderSW: https://github.com/YuichiSemura/CCFinderSW　Clone Notifier: https://ja.osdn.net/projects/clonenotifier/　CCEvovis: https://github.com/hirotaka0616/CCEvovis

[1] Y. Semura, N. Yoshida, E. Choi, and K. Inoue, “Multilingual Detection of Code Clones Using ANTLR Grammar Definitions,” in Proc. APSEC, December 2018, pp. 673-677
[2] 山中裕樹，崔恩瀞，吉田則裕，井上克郎，佐野建樹，“コードクローン変更管理システムの開発と実プロジェクトへの適用，”情報処理学会論文誌，vol.54, no.2, pp.883–893, 2013, 6.
[3] H. Honda, S. Tokui, K. Yokoi, E. Choi, N. Yoshida, and K. Inoue, “CCEvovis: A Clone Evolution Visualization System for Software Maintenance,” in Proc. ICPC, May 2019, pp. 122-125