

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7589109号
(P7589109)

(45)発行日 令和6年11月25日(2024.11.25)

(24)登録日 令和6年11月15日(2024.11.15)

(51)国際特許分類

G 0 6 F 8/77 (2018.01)

F I

G 0 6 F 8/77

請求項の数 7 (全28頁)

(21)出願番号 特願2021-94819(P2021-94819)
 (22)出願日 令和3年6月4日(2021.6.4)
 (65)公開番号 特開2022-186541(P2022-186541)
 A)
 (43)公開日 令和4年12月15日(2022.12.15)
 審査請求日 令和6年2月8日(2024.2.8)

(73)特許権者 000005108
 株式会社日立製作所
 東京都千代田区丸の内一丁目6番6号
 (74)代理人 110001689
 青稜弁理士法人
 市井 誠
 東京都千代田区丸の内一丁目6番6号
 株式会社日立製作所内
 川上 真澄
 東京都千代田区丸の内一丁目6番6号
 株式会社日立製作所内
 審査官 円子 英紀

最終頁に続く

(54)【発明の名称】 ソースコード解析装置およびソースコード解析方法

(57)【特許請求の範囲】

【請求項1】

ソフトウェア開発におけるアンチパターンまたはリファクタリングパターンを要因パターンとして、前記要因パターンがソフトウェア開発に対して与える影響を提示するソースコード解析装置であって、

前記ソースコード解析装置のユーザに情報を提示するユーザインターフェースと、

解析対象となるソフトウェアに関するチケット、ソースコード変更履歴およびソースコードから要因パターンの情報を取得する要因パターン取得部と、

前記要因パターンから要因メトリクスを算出する要因メトリクス算出部と、

解析対象となるチケット、変更履歴、ソースコードから品質、費用、納期のいずれかに関する計量値であるQCD計測値を取得するQCD計測値取得部と、

前記QCD計測値から品質、費用、納期を評価するQCD指標を算出するQCD指標算出部と、

前記要因メトリクスと前記QCD指標の相関を算出し、相関リストを生成する相関算出部と、

重要要因パターン分析部とを備え、

前記要因パターンとそれに関する要因メトリクスの重みづけに関する情報と、

前記QCD計測値とそれに関するQCD指標の重みづけに関する情報とを保持し、

前記チケット、前記ソースコード変更履歴、前記ソースコードのいずれかまたはその組合せが入力されると、

10

20

前記要因パターン取得部は、入力された前記チケット、前記ソースコード変更履歴、前記ソースコードから要因パターンの情報を生成し、

前記要因メトリクス算出部は、前記要因パターンに関する計量情報として、前記要因パターンに関する要因メトリクスを算出し、

前記QCD計測値取得部は、入力された前記チケット、前記ソースコード変更履歴、前記ソースコードから前記QCD計測値を取得し、

前記QCD指標算出部は、前記QCD計測値に基づき、前記QCD指標を算出し、

前記相関算出部は、前記要因メトリクスと前記QCD指標のペアに対する相関係数を算出し、前記相関リストに格納し、

前記重要要因パターン分析部は、前記要因パターンと前記QCD計測値のペアに対して、
前記相関リストと、前記要因パターンとそれに関する要因メトリクスの重みづけに関する情報と、前記QCD計測値とそれに関するQCD指標の重みづけに関する情報とに基づき、スコアを算出し、前記要因パターンと前記QCD計測値のペアに対してスコアを付けた重要要因パターン情報テーブルを生成し、

前記ユーザインタフェースは、ユーザに前記重要要因パターン情報テーブルの情報を提示することを特徴とするソースコード解析装置。

【請求項2】

前記重要要因パターン分析部は、前記要因パターンと前記QCD計測値のペアに対して、順位を算出し、前記要因パターンと前記QCD計測値のペアの順位を前記重要要因パターン情報テーブルに格納することを特徴する請求項1記載のソースコード解析装置。

【請求項3】

前記重要要因パターン分析部は、前記要因パターンと前記QCD計測値のペアに対して、前記QCD計測値が、品質、費用、納期のいずれによるものであるかの情報を前記重要要因パターン情報テーブルに格納することを特徴する請求項1記載のソースコード解析装置。

【請求項4】

さらに、前記要因パターンの情報に要因パターンが計量された要因パターンの発生時点の情報を付加する要因メトリクス時系列情報付加部と、

前記QCD指標の情報に前記QCD指標の算出の元となったQCD計測値に関する事象の発生時点の情報を付加する要因パターン指標時系列情報付加部と、

要因パターンが計量された要因パターンの発生時点の情報が付加された要因メトリクスの情報と、前記QCD指標の算出の元となったQCD計測値に関する事象の発生時点の情報が付加されたQCD指標の情報から、要因メトリクスを説明変数、QCD指標を目的変数とする遅延相関分析を行い、遅延相関分析の係数を保持する遅延相関リストを生成する遅延相関算出部とを備え、

前記遅延相関算出部は、前記要因メトリクスと前記QCD指標のペアに対する遅延相関係数を算出し、前記遅延相関リストに格納することを特徴とする請求項1記載のソースコード解析装置。

【請求項5】

前記遅延相関算出部は、要因メトリクスとQCD指標に関する影響期間を算出し、前記遅延相関リストに格納することを特徴とする請求項4記載のソースコード解析装置。

【請求項6】

ソフトウェア開発におけるアンチパターンまたはリファクタリングパターンを要因パターンとして、前記要因パターンがソフトウェア開発に対して与える影響を提示するソースコード解析装置によるソースコード解析方法であって、

前記ソースコード解析装置は、

前記ソースコード解析装置のユーザに情報を提示するユーザインタフェースと、

解析対象となるソフトウェアに関するチケット、ソースコード変更履歴およびソースコードから要因パターンの情報を取得する要因パターン取得部と、

前記要因パターンから要因メトリクスを算出する要因メトリクス算出部と、

10

20

30

40

50

解析対象となるチケットおよび変更履歴およびソースコードから品質、費用、納期のいずれかに関する計量値である Q C D 計測値を取得する Q C D 計測値取得部と、

前記 Q C D 計測値から品質、費用、納期を評価する Q C D 指標を算出する Q C D 指標算出部と、

前記要因メトリクスと前記 Q C D 指標の相関を算出し、相関リストを生成する相関算出部と、

重要要因パターン分析部とを備え、

前記要因パターンとそれに関する要因メトリクスの重みづけに関する情報と、

前記 Q C D 計測値とそれに関する Q C D 指標の重みづけに関する情報とを保持し、

前記ソースコード解析装置が、前記チケット、前記ソースコード変更履歴、前記ソースコードのいずれかまたはその組合せの入力を受け付けるステップ、

前記要因パターン取得部が、入力された前記チケット、前記ソースコード変更履歴、前記ソースコードから要因パターンの情報を生成するステップと、

前記要因メトリクス算出部が、前記要因パターンに関する計量情報として、前記要因パターンに関する要因メトリクスを算出するステップと、

前記 Q C D 計測値取得部が、入力された前記チケット、前記ソースコード変更履歴、前記ソースコードから前記 Q C D 計測値を取得するステップと、

前記 Q C D 指標算出部が、前記 Q C D 計測値に基づき、前記 Q C D 指標を算出するステップと、

前記相関算出部が、前記要因メトリクスと前記 Q C D 指標のペアに対する相関係数を算出し、前記相関リストに格納するステップと、

前記重要要因パターン分析部が、前記要因パターンと前記 Q C D 計測値のペアに対して、前記相関リストと、前記要因パターンとそれに関する要因メトリクスの重みづけに関する情報と、前記 Q C D 計測値とそれに関する Q C D 指標の重みづけに関する情報に基づき、スコアを算出し、前記要因パターンと前記 Q C D 計測値のペアに対してスコアを付けた重要要因パターン情報テーブルを生成するステップと、

前記ユーザインタフェースにより、ユーザに前記重要要因パターン情報テーブルの情報を提示するステップとを有することを特徴とするソースコード解析方法。

【請求項 7】

前記ソースコード解析装置は、

さらに、前記要因パターンの情報に要因パターンが計量された要因パターンの発生時点の情報を付加する要因メトリクス時系列情報付加部と、

前記 Q C D 指標の情報に前記 Q C D 指標の算出の元となった Q C D 計測値に関する事象の発生時点の情報を付加する要因パターン指標時系列情報付加部と、

要因パターンが計量された要因パターンの発生時点の情報が付加された要因メトリクスの情報と、前記 Q C D 指標の算出の元となった Q C D 計測値に関する事象の発生時点の情報が付加された Q C D 指標の情報から、要因メトリクスを説明変数、 Q C D 指標を目的変数とする遅延相関分析を行い、遅延相関分析の係数を保持する遅延相関リストを生成する遅延相関算出部とを備え、

前記遅延相関算出部が、前記要因メトリクスと前記 Q C D 指標のペアに対する遅延相関係数を算出し、前記遅延相関リストに格納するステップを有することを特徴とする請求項 6記載のソースコード解析方法。

【発明の詳細な説明】

【技術分野】

【 0 0 0 1 】

本発明は、ソースコード解析装置およびソースコード解析方法に係り、ソフトウェア開発のリファクタリングにおけるソースコードを解析するにあたり、ソフトウェア開発上の問題点の発見を容易にし、ソフトウェア品質向上とソフトウェアの開発効率を上げるのに好適なソースコード解析装置およびソースコード解析方法に関する。

【背景技術】

10

20

30

40

50

【0002】

近年のソフトウェア開発では、開発済みの母体ソフトウェアを拡張または変更することによって新しいソフトウェアを開発する派生開発が主流である。ソフトウェアの派生開発においては、長年に渡る機能拡張または変更の繰り返しにより、ソフトウェアが複雑化し、ソースコードの可読性が低下しやすい。このような状況を解決するため、ソフトウェアのリファクタリングが一般的に行われている。リファクタリングとは、ソフトウェアの振る舞いを変えず、その内部構造を変更することである。保守性が低いメソッド、クラス、ファイルなどのプログラムの構成要素（以下、「プログラム要素」という）を適切にリファクタリングすれば、ソフトウェアが拡張または変更しやすくなる。

【0003】

リファクタリングの際によく行われる手法の一つとしては、アンチパターンに着目することがある。アンチパターンとは、保守性が低いプログラム要素が有する、ソフトウェア開発上において問題とされる特徴（以下、「問題特徴」という）と、当該問題特徴を持つプログラム要素のリファクタリング方法のパターンをまとめたものである。リファクタリング対象ソフトウェアにおいて、アンチパターンが示すような問題特徴を有するプログラム要素を特定できれば、アンチパターンを改善することにより、ソフトウェアのリファクタリングを行うことができる。

【0004】

ソフトウェアのソースコード解析にあたり、アンチパターンを見出す技術については、例えば、特許文献1に開示がある。特許文献1には、ソフトウェアのソースコード解析にあたり、ソースコード、プログラム要素変更履歴からロジカルカップリンググラフを生成し、ソースコードのアンチパターン該当可否を評価する手法が記載されている。また、特許文献2には、ソースコード、プログラム要素変更履歴からメトリクスを計測し、ソースコードのアンチパターン該当可否を評価する手法が記載されている。これらのことにより、リファクタリングが必要ない複雑なソースコードを抽出せず、真にリファクタリングが必要なソースコードを特定することができる。

【0005】

また、非特許文献1には、OSS（Open Source Software）のソフトウェア開発過程を分析するために、時間的順序を考慮した相関分析（遅延相関分析）を行う技術が開示されている。

【先行技術文献】

【特許文献】

【0006】

【文献】特開2019-79312号公報

【文献】特開2019-219848号公報

【非特許文献】

【0007】

【文献】山谷ら，“OSSシステムとコミュニティの共進化の理解を目的としたデータマイニング手法”，情報処理学会論文誌，Vol.56，No.1，2015

【発明の概要】

【発明が解決しようとする課題】

【0008】

上記従来技術の特許文献1、および、特許文献2に記載されたソースコードの解析手法は、いずれも特定種類のアンチパターンをソースコード中から検出するものである。そのため、ソフトウェアの保守要因は、アンチパターン検出に先立って、どのようなアンチパターンを抽出するかを判断する必要があるが、どのような種類のアンチパターンを優先的に改善すべきかは、ソフトウェアの性質等によって異なるため、リファクタリングを行うソフトウェアの保守要因に高度な知見が必要となる。例えば、特許文献1の方法により検出されるアンチパターンとして、Shotgun surgeryがある。Shotgun surgeryは、当該プログラム要素を拡張または変更すると多くの他のプログラム要

10

20

30

40

50

素も同時に拡張または変更されるという問題特徴である。また、例えば、特許文献2の方法により検出されるアンチパターンとして、`Long method`がある。`Long method`は、プログラム要素としてメソッドのコード数が大きすぎ、プログラム要素の可読性が低くなるという問題特徴である。このように、特許文献1、特許文献2に記載された手法では、それぞれ異なる観点から特定のソースコードのパターンにおける問題点を指摘することはできるが、特許文献1の方法で指摘された問題点と、特許文献2の方法で指摘された問題点のいずれがよりソフトウェア開発上の障壁となる問題であるかは判定することができない。

【0009】

さらに、リファクタリングの際には、リファクタリングパターンとよばれる変更方法によりソースコードの改善が行われる。例えば、リファクタリングパターンの一例として、あるメソッドを別のメソッドへ移動する`Move method`がある。また、リファクタリングパターンの別の一例として、メソッドの一部を別の新しいメソッドとして切り出す`Extract method`がある。しかしながら、これらのように、リファクタリングパターンによってソースコードを変更する効果は、アンチパターンの悪影響と同様、ソフトウェアの性質等によって異なるため、リファクタリングを行うソフトウェアの保守要員に高度な知見が必要となる。

【0010】

本発明の目的は、ソフトウェア開発のリファクタリングにおけるソースコードを解析するにあたり、異なる種類のアンチパターン検出結果、あるいは、リファクタリングパターン適用結果を定量的に評価し、ソフトウェア開発上の問題点の発見を容易にし、ソフトウェア品質向上とソフトウェアの開発効率を上げることのできるソースコード解析装置およびソースコード解析方法を提供することにある。

【課題を解決するための手段】

【0011】

本発明のソースコード解析装置の構成は、好ましくは、ソフトウェア開発におけるアンチパターンまたはリファクタリングパターンを要因パターンとして、要因パターンがソフトウェア開発に対して与える影響を提示するソースコード解析装置であって、解析対象となるソフトウェアに関するチケット、ソースコード変更履歴およびソースコードから要因パターンの情報を取得する要因パターン取得部と、要因パターンから要因メトリクスを算出する要因メトリクス算出部と、解析対象となるチケット、変更履歴、ソースコードから品質、費用、納期のいずれかに関する計量値であるQCD計測値を取得するQCD計測値取得部と、QCD計測値から品質、費用、納期を評価するQCD指標を算出するQCD指標算出部と、要因メトリクスとQCD指標の相関を算出し、相関リストを生成する相関算出部とを備え、チケット、ソースコード変更履歴、ソースコードのいずれかまたはその組合せが入力されると、要因パターン取得部は、入力されたチケット、ソースコード変更履歴、ソースコードから要因パターンの情報を生成し、要因メトリクス算出部は、要因パターンに関する計量情報として、要因パターンに関する要因メトリクスを算出し、QCD計測値取得部は、入力されたチケット、ソースコード変更履歴、ソースコードからQCD計測値を取得し、QCD指標算出部は、QCD計測値に基づき、QCD指標を算出し、相関算出部は、要因メトリクスとQCD指標のペアに対する相関係数を算出し、相関リストに格納するようにしたものである。

【発明の効果】

【0012】

本発明によれば、ソフトウェア開発のリファクタリングにおけるソースコードを解析するにあたり、異なる種類のアンチパターン検出結果、あるいは、リファクタリングパターン適用結果を定量的に評価し、ソフトウェア開発上の問題点の発見を容易にし、ソフトウェア品質向上とソフトウェアの開発効率を上げることのできるソースコード解析装置およびソースコード解析方法を提供することができる。

【図面の簡単な説明】

10

20

30

40

50

【0013】

- 【図1】実施形態1に係るソースコード解析装置の機能構成図である。
- 【図2】ソースコード解析装置のハードウェア・ソフトウェア構成図である。
- 【図3】実施形態1に係るソースコード解析装置におけるデータフロー図である。
- 【図4】要因パターン属性テーブルの一例を示す図である。
- 【図5】要因メトリクステーブルの一例を示す図である。
- 【図6】QCD計測値属性テーブルの一例を示す図である。
- 【図7】QCD指標テーブルの一例を示す図である。
- 【図8】相関リストの一例を示す図である。
- 【図9】要因パターン・要因メトリクス定義テーブルの一例を示す図である。
- 【図10】QCD計測値・QCD指標定義テーブルの一例を示す図である。
- 【図11】重要要因パターン情報テーブルの一例を示す図である。
- 【図12】要因パターン取得処理を示すフローチャートである。
- 【図13】要因メトリクス算出処理を示すフローチャートである。
- 【図14】QCD計測値取得処理を示すフローチャートである。
- 【図15】QCD指標算出処理を示すフローチャートである。
- 【図16】相関算出処理を示すフローチャートである。
- 【図17】重要要因パターン分析処理を示すフローチャートである。
- 【図18】重要要因パターン評価画面の一例を示す図である。
- 【図19】遅延相関分析の概念について説明する図である。

10

- 【図20】実施形態2に係るソースコード解析装置の機能構成図である。
- 【図21】実施形態2に係るソースコード解析装置におけるデータフロー図である。
- 【図22】時系列要因メトリクステーブルの一例を示す図である。
- 【図23】時系列QCD指標テーブルの一例を示す図である。
- 【図24】遅延相関リストの一例を示す図である。
- 【図25】遅延相関分析による重要要因パターン情報テーブルの一例を示す図である。
- 【図26】要因メトリクス時系列情報付加処理を示すフローチャートである。
- 【図27】QCD指標時系列情報付加処理を示すフローチャートである。
- 【図28】遅延相関算出処理を示すフローチャートである。
- 【図29】遅延相関係数算出処理を示すフローチャートである。

20

- 【図30】遅延相関分析による重要パターン分析処理を示すフローチャートである。

30

【発明を実施するための形態】

【0014】

以下、本発明に係る各実施形態を、図1ないし図30を用いて説明する。

【0015】

〔実施形態1〕

以下、本発明に係る実施形態1を、図1ないし図18を用いて説明する。

【0016】

本実施形態は、ソフトウェア開発のリファクタリングにおけるソースコードを解析するにあたり、異なる種類のアンチパターン検出結果、あるいは、リファクタリングパターン適用結果を定量的に評価するソースコード解析装置に関するものである。以下、アンチパターンとリファクタリングパターンをまとめて、「要因パターン」ということにする。

40

【0017】

また、本実施形態において、メソッドもしくは関数と呼ばれるプログラム要素を対象とするが、そのプログラム要素の単位は、メソッドに限らず、クラスやモジュール構造体、構造体など、要因パターンで定義される単位であればどのような単位でも適用可能である。

【0018】

先ず、図1ないし図3を用いてソースコード解析装置の構成とソースコード解析装置が実行する処理の概要について説明する。

【0019】

50

ソースコード解析装置 10 は、機能部として、図 1 に示されるように、要因パターン取得部 101、要因メトリクス算出部 102、QCD 計測値取得部 103、QCD 指標算出部 104、相関算出部 105、重要要因パターン分析部 106、情報記憶部 120 を備える。

【0020】

要因パターン取得部 101 は、解析対象となるチケット 121、ソースコード変更履歴 122、ソースコード 123 を受け付けて、要因パターン属性テーブル 124 を生成する機能部である。

【0021】

要因メトリクス算出部 102 は、要因パターン属性テーブル 124 の情報を受付けて、要因メトリクステーブル 125 を生成する機能部である。

10

【0022】

QCD 計測値取得部 103 は、チケット 121、ソースコード変更履歴 122、ソースコード 123 を受け付けて、QCD 計測値テーブル 126 を生成する機能部である。

【0023】

QCD 指標算出部 104 は、QCD 計測値テーブル 126 の情報を受け付けて、QCD 指標テーブル 127 を生成する機能部である。

【0024】

相関算出部 105 は、要因メトリクステーブル 125 と、QCD 指標テーブル 127 との情報を受け付けて、相関リスト 128 を生成する機能部である。

20

【0025】

重要要因パターン分析部 106 は、相関リスト 128、要因パターン・要因メトリクス定義テーブル 129、QCD 計測値・QCD 指標定義テーブル 130 の情報を受け付けて、重要要因パターン情報テーブル 131 を生成する機能部である。

【0026】

情報記憶部 120 は、ソースコード解析装置 10 に使用されるデータを記憶する機能部である。ソースコード解析装置 10 の情報記憶部 120 には、チケット 121、ソースコード変更履歴 122、ソースコード 123、要因パターン属性テーブル 124、要因メトリクステーブル 125、QCD 計測値テーブル 126、QCD 指標テーブル 127、相関リスト 128、要因パターン・要因メトリクス定義テーブル 129、QCD 計測値・QCD 指標定義テーブル 130、重要要因パターン情報テーブル 131 が記憶される。情報記憶部 120 は、これらの情報以外にも、要因パターン取得部 101、要因メトリクス算出部 102、QCD 計測値取得部 103、QCD 指標算出部 104、相関算出部 105、重要要因パターン分析部 106 が適宜参照あるいは生成する情報を記憶する。情報記憶部 120 は、補助記憶装置 16 に記憶する情報を、例えば、ファイルシステムや DBMS (DataBase Management System) によって管理する。

30

なお、ソースコード解析装置 10 で使用されるテーブルの詳細については、後に説明する。

【0027】

次に、図 2 を用いてソースコード解析装置のハードウェア・ソフトウェア構成について説明する。

40

ソースコード解析装置 10 は、ソフトウェアシステムの開発や保守等に利用される情報処理装置であり、一般的な PC (Personal Computer) で実現することができる。また、ソースコード解析装置 10 は、クラウドシステムにより提供されるクラウドサーバ等のように仮想的に実現されるものであってもよい。

【0028】

ソースコード解析装置 10 は、図 2 に示されるように、例えば、プロセッサ 11、主記憶装置 12、通信装置 13、入力装置 14、出力装置 15、および補助記憶装置 16 を備える。これらは、コンポーネントは、バス等の通信手段を介して互いに通信できるように接続されている。

50

【0029】

プロセッサ 11 は、例えば、C P U (Central Processing Unit) や M P U (Micro Processing Unit) のような半導体装置によりロジック回路を実現したものである。プロセッサ 11 が、主記憶装置 12 に格納されているプログラムを読み出して実行することにより、ソースコード解析装置 10 の様々な機能が実現される。

【0030】

主記憶装置 12 は、プログラムやデータを一時的に記憶する半導体装置であり、例えば、R A M (Random Access Memory)、不揮発性半導体メモリ (N V R A M (Non Volatile RAM)) 等である。

【0031】

通信装置 13 は、L A N やインターネット等の通信手段を介した他の装置との間の通信を実現する有線、または、無線による通信インターフェースを実現する装置であり、例えば、N I C (Network Interface Card)、無線通信モジュール、U S B (Universal Serial Interface) モジュール、シリアル通信モジュール等である。

10

【0032】

入力装置 14 は、ソースコード解析装置 10 に情報を入力する装置であり、例えば、キーボード、マウス、タッチパネル、カードリーダ、音声入力装置等である。出力装置 15 は、ソースコード解析装置 10 における処理経過や処理結果等の各種情報をユーザに提供するユーザインターフェースを実現する装置であり、例えば、画面表示装置 (液晶モニタ、L C D (Liquid Crystal Display)、グラフィックカード等)、音声出力装置 (スピーカ等)、印字装置等である。なお、例えば、ソースコード解析装置 10 が通信装置 13 を介して他の装置との間で情報の入力や出力を構成としてもよい。

20

【0033】

補助記憶装置 16 は、例えば、H D D (Hard Disk Drive)、S S D (Solid State Drive)、光学式記憶装置 (C D (Compact Disc)、D V D (Digital Versatile Disc) 等)、ストレージシステム、I C カード、S D メモリカードや光学式記録媒体等の記録媒体の読み取り / 書込装置、クラウドサーバの記憶領域等である。補助記憶装置 16 に格納されているプログラムやデータは主記憶装置 12 に随時ロードされる。

【0034】

本実施形態のソースコード解析装置 10 の補助記憶装置 16 には、要因パターン取得プログラム 161、要因メトリクス算出プログラム 162、Q C D 計測値取得プログラム 163、Q C D 指標算出プログラム 164、相關算出プログラム 165、重要要因パターン分析プログラム 166 がインストールされている。

30

【0035】

要因パターン取得プログラム 161、要因メトリクス算出プログラム 162、Q C D 計測値取得プログラム 163、Q C D 指標算出プログラム 164、相關算出プログラム 165、重要要因パターン分析プログラム 166 は、各々要因パターン取得部 101、要因メトリクス算出部 102、Q C D 計測値取得部 103、Q C D 指標算出部 104、相關算出部 105、重要要因パターン分析部 106 の機能を実現するプログラムである。

【0036】

また、図 2 には図示しなかったが、本実施形態のソースコード解析装置 10 の補助記憶装置 16 には、チケット 121、ソースコード変更履歴 122、ソースコード 123、要因パターン属性テーブル 124、要因メトリクステーブル 125、Q C D 計測値テーブル 126、Q C D 指標テーブル 127、相關リスト 128、要因パターン・要因メトリクス定義テーブル 129、Q C D 計測値・Q C D 指標定義テーブル 130、重要要因パターン情報テーブル 131 が格納されている。

40

【0037】

本実施形態では、補助記憶装置 16 にインストールされたプログラムを、ロードしてプロセッサ 11 が、それを実行することにより、その機能を実現する例を説明したが、これらの機能は、例えば、ソースコード解析装置 10 が備えるハードウェア (A S I C (AppI

50

ication Specific Integrated Circuit) 等) によって実現されてもよい。

【0038】

次に、図3を用いて実施形態1に係るソースコード解析装置10で実行される処理の概要とデータフローについて説明する。

先ず、図3に示されるように、要因パターン取得部101は、チケット121と、ソースコード変更履歴122と、ソースコード123とを受け付けて、要因パターン属性テーブル124を生成する。チケット121とは、チケット開発駆動という開発手法で提唱された概念に基づくものであり、ソフトウェア開発におけるなんらかの課題、ソフトウェア開発者のコミットメントを提示するためのデータである。ソースコード変更履歴122は、ソフトウェア開発のターゲットであるソースコードを新規に開発した、あるいは、それを変更した履歴のデータである。要因パターン属性テーブル124とは、要因パターンごとの様々な属性を格納するテーブルである。

【0039】

次に、要因メトリクス算出部102は、要因パターン属性テーブル124の情報を受付けて、要因メトリクステーブル125を生成する。要因メトリクステーブル125とは、プログラム要素ごとに、要因パターン属性テーブル124に格納されている要因パターンの属性に基づいたソフトウェア評価のための計量値を要因メトリクスとして格納するテーブルである。

【0040】

次に、QCD計測値取得部103は、チケット121、ソースコード変更履歴122、ソースコード123を受け付けて、QCD計測値テーブル126を生成する。QCD計測値テーブル126は、プログラムのQCD (Quality, Cost, Delivery: 品質、費用、納期) の観点から評価するQCD計測値を格納するためのテーブルである。

【0041】

次に、QCD指標算出部104は、QCD計測値テーブル126の情報を受け付けて、QCD指標テーブル127を生成する。QCD指標テーブル127は、プログラム要素ごとに、QCD計測値テーブル126に格納されているQCD計測値に基づいた指標を格納するテーブルである。

【0042】

次に、相関算出部105は、要因メトリクステーブル125と、QCD指標テーブル127との情報を受け付けて、要因メトリクステーブル125に格納された要因メトリクスとQCD指標テーブル127に格納されたQCD指標との相関係数を算出し、相関リスト128を生成する。相関リスト128は、要因メトリクステーブル125に格納された要因メトリクスとQCD指標テーブル127に格納されたQCD指標との相関係数を格納するリストである。

【0043】

次に、重要要因パターン分析部106は、相関リスト128、要因パターン・要因メトリクス定義テーブル129、QCD計測値・QCD指標定義テーブル130の情報を受け付けて、重要要因パターンを評価するスコアとそのスコアに基づいた要因パターンの重要性の順位を算出し、重要要因パターン情報テーブル131を生成する。要因パターン・要因メトリクス定義テーブル129は、要因パターンの種別とそれに対する要因メトリクスを評価する重みを定義するためのテーブルである。QCD計測値・QCD指標定義テーブル130は、QCD計測値とそれに対するQCD指標を評する重みを定義するためのテーブルである。重要要因パターン情報テーブル131は、要因パターンの種別、QCD計測値の種別ごとに、重要要因パターンを評価するスコアとそのスコアに基づいた要因パターンの重要性の順位を格納するテーブルである。

【0044】

次に、図4ないし図11を用いて実施形態1のソースコード解析装置で使用されるデータ構造について説明する。

【0045】

10

20

30

40

50

要因パターン属性テーブル 124 は、要因パターンごとの様々な属性を格納するテーブルであり、ソースコードのチケット 121、ソースコード変更履歴 122、ソースコード 123 を入力し、要因パターン取得部 101 によって算出された様々な要因パターンに関する情報を格納するテーブルである。

【0046】

要因パターン属性テーブル 124 は、図 4 に示されるように、要因パターン名 124a、検出 ID 124b、プログラム要素 124c、属性値 i_{124d_i} ($i = 1, 2, \dots$) の各フィールドからなる。

【0047】

要因パターン名 124a には、当該レコードが対象とする要因パターンの名称が格納される。検出 ID 124b には、検出箇所を一意に特定するための識別子が格納される。検出箇所は、例えば、あるファイルに記述されたプログラムの何行目などとして記録される。プログラム要素 124c には、検出箇所となるプログラム要素を識別する名称または識別子が格納される。属性値 i_{124d_i} には、要因パターン毎に捉えられる様々な属性の値が格納される。

10

【0048】

要因パターンの一例としては、特許文献 1 に示されるアンチパターン「Shotgun Surgery」（一つのプログラム上の変更を複数クラス、複数の関数、メソッドに対して同時に行われる）、「Long method」（長すぎるメソッド）、「Magic numbers」（説明のない数値をアルゴリズム上に埋め込む）などがある。

20

【0049】

要因パターンの属性値は、関連するプログラム要素数や、例えば、「Long method」の場合は、メソッドの記述したコード数、「Magic numbers」の場合には、説明のない数値がプログラムに出現した回数などである。また、適合する要因パターンの場合には、バスケット解析ルール評価値である支持度、確信度、リフト値などを属性値として採用することもできる。

【0050】

要因メトリクステーブル 125 は、プログラム要素ごとに、要因パターン属性テーブル 124 に格納されている要因パターンの属性に基づいたソフトウェア評価のため計量値を要因メトリクスとして格納するテーブルであり、要因パターン属性テーブル 124 を入力とし、要因メトリクス算出部 102 によってプログラム要素単位で算出された情報を格納するテーブルである。

30

【0051】

要因メトリクステーブル 125 は、図 5 に示されるように、プログラム要素名 125a、要因メトリクス i_{125b_i} ($i = 1, 2, \dots$) のフィールドからなる。

【0052】

プログラム要素名 125a には、当該レコードのプログラム要素を一意に特定するプログラム要素の名称が格納される。要因メトリクス i_{125b_i} には、対象となるプログラム要素の要因パターンに関する計量値としてメトリクスの値が格納される。要因メトリクスは、一つまたは複数の要因パターンの属性値から算出する値である。

40

【0053】

要因メトリクスは、例えば、「Long method」の場合は、コメントのコード数を除外する、または、ファイルごとの行数の違いを考慮して、一行の文字数を一定として全ての行数を比較可能ないようにした値（例えば、0 ~ 1 の値に正規化など）である。

【0054】

QCD 計測値テーブル 126 は、プログラムの QCD の観点から評価する QCD 計測値に関する属性を格納するためのテーブルであり、解析対象となるソースコードのソフトウェア開発におけるチケット 121、ソースコード変更履歴 122、ソースコード 123 を入力し、QCD 計測値取得部 103 によって算出された様々な品質、コスト、納期に関する計測値に関する属性を格納する。

50

【0055】

QCD計測値テーブル126は、QCD計測値種別126a、計測ID126b、プログラム要素126c、計測値i126d_i(i=1, 2, ...)の各フィールドからなる。

【0056】

QCD計測値種別126aには、当該レコードのQCD計測値種別として名づけられた名称が格納される。計測ID126bには、当該レコードにQCD計測値種別に関する計測対象を一意に特定するための識別子が格納される。プログラム要素126cには、計測の対象となったプログラム要素を一意に識別する名称または識別子が格納される。計測値i126d_iには、当該レコードにQCD計測値種別に関するQCD計測値が格納される。

【0057】

QCD計測値は、例えば、チケットに関連づけられたプログラム要素のチケットのクローズ日時から起票日時までの期間として計測されるリードタイムである。また、例えば、チケットに関連づけられたプログラム要素に関して発生した不良数である。

【0058】

QCD指標テーブル127は、プログラム要素ごとに、QCD計測値テーブル126に格納されているQCD計測値に基づいた指標をQCD指標として格納するテーブルであり、QCD計測値テーブル126を入力とし、QCD指標算出部104によってプログラム要素単位で算出された情報が格納される。

【0059】

QCD指標テーブル127は、図7に示されるように、プログラム要素名127a、QCD指標i127b_i(i=1, 2, ...)のフィールドからなる。

【0060】

プログラム要素名127aには、当該レコードのプログラム要素を一意に特定するプログラム要素の名称が格納される。QCD指標i127b_iには、対象となるプログラム要素に対するQCD指標の値が格納される。QCD指標は、一つまたは複数のQCD計測値から算出される値であり、例えば、複数のQCD計測値の代表値から算出される値を正規化した値となる。

【0061】

相関リスト128は、要因メトリクステーブル125に格納された要因メトリクスとQCD指標テーブル127に格納されたQCD指標との相関係数を格納するリストであり、要因メトリクステーブル125と、QCD指標テーブル127とを入力とし、相関算出部105によって算出された情報を格納する。

【0062】

相関リスト128は、図8に示されるように、要因メトリクス名128a、QCD指標名128b、相関係数128cの各フィールドからなる。

【0063】

要因メトリクス名128a、QCD指標名128bは、それぞれ相関分析の対象となる要因メトリクスの名称とQCD指標の名称が格納される。相関係数128cには、要因メトリクス名128aの要因メトリクスと、QCD指標名128bのQCD指標に対して相関分析を行って得られる相関係数の値が格納される。

【0064】

要因パターン・要因メトリクス定義テーブル129は、要因パターンの種別とそれに対する要因メトリクスを評価する重みを定義するためのテーブルである。

【0065】

要因パターン・要因メトリクス定義テーブル129は、図9に示されるように、要因パターン種別名129a、要因メトリクス名129b、重み129cの各フィールドからなる。

【0066】

要因パターン種別名129a、要因メトリクス名129bには、それぞれ重みを与える要因パターン種別の名称と、要因メトリクスの名称が格納される。重み129cには、要

10

20

30

40

50

因パターン種別ごとに、要因メトリクスをどのように重みづけて評価するかを定義する値が格納する。例えば、要因パターン種別名 129a に格納された同じ要因パターン種別の名称を有する重み 129c の値は、それらを合計して 1.0 になるような正の実数値が格納される。

【0067】

QCD 計測値・QCD 指標定義テーブル 130 は、QCD 計測値とそれに対する QCD 指標を評価するための重みを定義するためのテーブルである。QCD 計測値・QCD 指標定義テーブル 130 は、図 10 に示されるように、QCD 計測値名 130a、QCD 指標名 130b、影響観点 130c、重み 130d の各フィールドからなる。

【0068】

QCD 計測値名 130a、QCD 指標名 130b には、それぞれ重みを評価する QCD 計測値を識別するための QCD 計測値の名称と QCD 指標を識別するため QCD 指標の名称が格納される。影響観点 130c には、QCD 指標の観点として、品質、コスト、納期の種別を示す情報が格納され、例えば、QCD 指標がソフトウェアの品質のある一面を示したものであれば「Q」、ソフトウェアにかかるコストのある一面を示したものであれば「C」、ソフトウェアのリリースに要する時間のある一面を示したものであれば「D」が格納される。

【0069】

重み 130d には、QCD 計測値名 130a に対する QCD 指標名 130b に格納された QCD 指標の比率が格納される。例えば、QCD 計測値名 130a で同じ QCD 計測値名を有する QCD 計測値の QCD 指標の重みの総計は、1.0 として配分される。

10

【0070】

重要要因パターン情報テーブル 131 は、要因パターンの種別、QCD 計測値の種別ごとに、重要要因パターンを評価するスコアとそのスコアに基づいた要因パターンの重要性の順位を格納するテーブルであり、相関リスト 128、要因パターン・要因メトリクス定義テーブル 129、QCD 計測値・QCD 指標定義テーブル 130 を入力し、重要要因パターン分析部 106 によって算出された情報を格納するテーブルである。

【0071】

重要要因パターン情報テーブル 131 は、図 11 に示されように、要因パターン種別名 131a、QCD 計測値名 131b、影響観点 131c、スコア 131d、順位 131e の各フィールドからなる。

20

【0072】

要因パターン種別名 131a、QCD 計測値名 131b には、それぞれスコアを算出する対象となる要因パターン種別の名称と QCD 計測値の名称が格納される。影響観点 131c には、QCD 計測値・QCD 指標定義テーブル 130 と同様に、QCD 指標の観点として、品質、コスト、納期の種別を示す情報が格納される。スコア 131d には、これまでに算出した要因パターン種別と QCD 計測値のペアに対する相関係数、定義されている要因パターン種別と要因メトリクスのペアに対する重み、定義されている QCD 計測値、QCD 指標のペアに対する重みに基づいて算出される要因パターン種別と QCD 計測値のペアに対するソフトウェア開発においての影響を評価するためのスコアが格納される。順位 131e には、スコア 131d の値から見た要因パターン種別と QCD 計測値のペアに対するソフトウェア開発においての影響の順位が格納される。

30

【0073】

次に、図 12 ないし図 17 を用いて実施形態 1 に係るソースコード解析装置が行う処理について説明する。

40

【0074】

先ず、図 12 を用いて要因パターン取得処理について説明する。

先ず、ソースコード解析装置 10 の要因パターン取得部 101 は、解析対象プログラムに関連づけられたチケット 121 を読み込む (S101)。

次に、解析対象プログラムのソースコード変更履歴 122 を読み込む (S102)。

50

次に、解析対象プログラムのソースコード123を読み込む(S103)。

次に、全ての要因パターン種別Eに対して(S104-S106)、読み込まれたチケット121、ソースコード変更履歴122、ソースコード123に基づき、要因パターンを取得し(S105)、要因パターン属性テーブル124に必要な情報を格納する。要因パターンの取得は、既存のソフトウェア解析ツールを利用することができる。

【0075】

次に、図13を用いて要因メトリクス算出処理を説明する。

ソースコード解析装置10の要因メトリクス算出部102は、全ての要因パターン種別Eに対して(S201-S203)、要因パターン属性テーブル124に格納された情報に基づき、要因メトリクスを算出する(S202)。

10

【0076】

次に、図14を用いてQCD計測値取得処理を説明する。

先ず、ソースコード解析装置10のQCD計測値取得部103は、解析対象プログラムに関連づけられたチケット121を読み込む(S301)。

次に、解析対象プログラムのソースコード変更履歴122を読み込む(S302)。

次に、解析対象プログラムのソースコード123を読み込む(S303)。

次に、全てのQCD計測種別Eに対して(S304-S306)、QCD計測値を取得し(S305)、図6に示されるQCD計測値テーブル126に情報を格納する。QCD計測値の取得は、既存のソフトウェア開発管理ツールを用いて行うことができる。

20

【0077】

次に、図15を用いてQCD指標算出処理を説明する。

ソースコード解析装置10のQCD指標算出部104は、QCD計測値テーブル126に格納されている全てのQCD計測値Eに対して(S401-S402)、QCD計測値テーブル126に格納された情報に基づいて、QCD指標を算出し(S402)、図7に示されるQCD指標テーブル127にQCD指標を格納する。

【0078】

次に、図16を用いて相関算出処理を説明する。

ソースコード解析装置10の相関算出部105は、全ての要因メトリクスE1に対して(S501-S505)、S502-S504の処理を実施する。

30

【0079】

S501-S505のループの中で、全てのQCD指標E2に対して(S502-S10)、要因メトリクスE1、QCD指標E2の相関係数を算出する(S503)。例えば、相関算出部105は、Spearmanの順位相関係数を用いて相関係数を算出する。Spearmanの順位相関係数は、二つの変量が順位データであるときに、相関を求める指標であり、以下のように評価する。元データとなる要因メトリクス、QCD指標のスコアを順位に変換し、各ペアにおける二つの変数の順位の差Dを算出する。このとき、Spearmanの順位相関係数は、以下の(式1)で与えられる。

【0080】

【数1】

40

$$\rho = 1 - \frac{6 \sum D^2}{N^3 - N} \quad \dots \quad (\text{式1})$$

【0081】

ここで、 ρ は、要因メトリクス、QCD指標に関するSpearmanの順位相関係数、Dは、要因メトリクス値とQCD指標の値の順位の差、Nは、値のペアの数であり、 ρ は、全て要因メトリクス値とQCD指標のペアの要因メトリクス値とQCD指標の値の順

50

位の差の和をすることを意味する。Spearmanの順位相関係数は、変量が正規分布などの分布に従わない場合でも適用できるという利点がある。

【0082】

また、相関算出部105は、例えば、入力された要因メトリクステーブル125に対し、ボックスプロットを用いた外れ値除去を行った上で、相関係数を算出する。また、相関算出部105は、例えば、入力されたQCD指標テーブル127に対し、ボックスプロットを用いた外れ値除去を行った上で、相関係数を算出する。また、相関算出部105は、例えば、入力された要因メトリクステーブル125に対し、(メトリクス値の最大値-メトリクス値の最小値)でメトリクス値を除算することにより、値の正規化を行った上で、相関係数を算出する。また、相関算出部105は、例えば、入力されたQCD指標テーブル127に対し、(QCD指標値の最大値-QCD指標値の最小値)でQCD指標値を除算することにより、値の正規化を行った上で、相関係数を算出する。

10

【0083】

次に、図17を用いて重要要因パターン分析処理を説明する。

ソースコード解析装置10の重要要因パターン分析部106は、全ての要因メトリクスE1、QCD計測値E2の組(E1, E2)に対して(S601-S604)、S602、S603を実行する。

【0084】

S601-S604のループの中で、要因メトリクスE1に対応する要因パターン・要因メトリクス定義テーブル129の要因パターン種別名および重みを取得し、要因パターンのスコアを、相関リスト128に記憶された相関係数と取得した重みに基づき算出する(S602)。要因パターンのスコアは、例えば、要因パターンの要因メトリクスの相関係数と重みの積により算出する。

20

【0085】

次に、QCD指標E2に対応するQCD計測値名および重みを、QCD計測値・QCD指標定義テーブル130より取得し、QCD指標のスコアを、相関リスト128に格納された相関係数と取得した重みに基づき算出する(S603)。QCD指標のスコアは、例えば、QCD指標の相関係数と重みの積により算出する。

【0086】

次に、全ての要因パターン・QCD計測値の組(E1, E2)に対して(S605-S608)、同じ要因パターンE1・QCD計測値E2を有する要因パターンのスコアとQCD指標のスコアの和を、組(E1, E2)について合算し(S605)、重要要因パターン情報テーブル131に格納する。合算方法は、例えば、合計をとる。また、例えば、正の値が良い影響、負の値が悪い影響を表し、最大値が100、最小値が-100となるように、値を加工する。

30

【0087】

要因パターン・QCD計測値の組(E1, E2)に対して、合算されたスコアの降順で並び替えることにより(すなわち、スコアが大きいものが順位が上)、順位を算出し(S607)、重要要因パターン情報テーブル131に格納する。

【0088】

次に、図18を用いてソースコード解析装置が提供するユーザインターフェースについて説明する。

40

図18は、重要要因パターン評価画面の一例を示す図である。

ソースコード解析装置10は、ソースコードの解析結果として、ディスプレイなど表示装置に、図18に示されるような重要要因パターン評価画面300を表示する。

重要要因パターン評価画面300は、開発プロジェクト表示領域310、重要要因パターン情報表示領域320を有する。

開発プロジェクト表示領域310には、関連するプロジェクトの情報が表示される。重要要因パターン情報表示領域320には、重要要因パターン情報テーブル131の情報が表示される。

50

【0089】

〔実施形態2〕

以下、本発明に係る実施形態2を、図19ないし図30を用いて説明する。

【0090】

実施形態1では、ソフトウェア開発におけるパターンと、ソフトウェア開発のQCDにおける計測値の相関分析を行い、ソフトウェア開発におけるパターンと、ソフトウェア開発のQCDにおける計測値のペアをスコア付けすることにより、ソフトウェア開発者に提示するソースコード解析装置の例について説明した。

【0091】

本実施形態では、実施形態1を前提として、遅延相関分析の手法も取り込み、実施形態1の相関係数の代わり、加算係数と、差分係数と、遅延係数によって表現される遅延相関係数によって、ソフトウェア開発におけるパターンと、ソフトウェア開発のQCDにおける計測値のペアをスコア付けするものである。

10

【0092】

以下、実施形態1と異なる所を中心に説明する。

【0093】

先ず、図19を用いて遅延相関分析の概念について説明する。

図19は、遅延相関分析の概念について説明する図である。

遅延相関分析とは、時間的順序を考慮した相関分析の手法であり、二つの時系列データを用意し、時間を遅延させながら、観察する関係を探る手法である。遅延相関分析では、一方の時系列データを目的変数とし、他方の時系列データを、説明変数とし、説明変数の一定期間の平均値と、目的変数の変化量の間の相関を求める。

20

【0094】

なお、遅延相関分析をソフトウェア開発に利用する例としては、非特許文献1に説明されている。

【0095】

今、図19に示されるように、 e_i を時刻iにおける説明変数とし、 r_n を時刻nにおける目的変数とする。

【0096】

遅延相関分析においては、説明変数の値が累積し、一定期間の遅延をもって目的変数の値の変化に影響を与えるということを表現するために、加算係数（説明変数の値を累積させた期間： $i - j$ ）、差分係数（目的変数の値の変化を考慮する期間： $n - m$ ）、遅延係数（説明変数が目的変数に影響を与えるまでの期間： $n - i$ ）の三つのパラメータを定義し、この三つのパラメータの最適値を、相関係数が最大となるよう求める処理を行う（詳細は、後述）。

30

【0097】

ある加算係数（ $i - j$ ）における説明変数の値の累積値を、 e_{ij} とするとき、 e_{ij} は、以下の（式2）で定義される。

【0098】

【数2】

40

$$e_{ij} = \frac{e_i + e_{i-1} + \dots + e_j}{i-j+1} \dots \quad (\text{式2})$$

【0099】

また、ある差分係数（ $n - m$ ）における目的変数の値の変化値を r_{nm} とするとき、 r_{nm} は、以下の（式3）で定義される。

50

【0100】

【数3】

$$r_{nm} = r_n - r_m$$

... (式3)

【0101】

遅延相関係数 c_{er} は、以下の(式4)で定義される。

【0102】

【数4】

$$C_{er} = \frac{S_{er}}{S_e S_r} \dots \text{(式4)}$$

【0103】

ここで、 S_e および S_r は、それぞれ e_{ij} , r_{nm} の標準偏差、 S_{er} は e_{ij} と r_{nm} の共分散である。

【0104】

このように、遅延相関分析は、説明変数の値が累積した結果、一定期間後に目的変数の値に変化を与えることを想定した相関分析手法である。

【0105】

次に、図20および図21を用いてソースコード解析装置の構成とソースコード解析装置が実行する処理の概要について説明する。

【0106】

図20は、実施形態2に係るソースコード解析装置の機能構成図である。

【0107】

図21は、実施形態2に係るソースコード解析装置におけるデータフロー図である。

【0108】

本実施形態のソースコード解析装置10は、機能部として、図20に示されるように、実施形態1のソースコード解析装置10と比較すると、相関算出部105がなくなり、要因メトリクス時系列情報付加部107、QCD指標時系列情報付加部108、遅延相関算出部109、遅延相関分析による重要要因パターン分析部110が付け加わっている。

【0109】

要因メトリクス時系列情報付加部107は、要因メトリクステーブル125の各々のレコードに要因メトリクスが計量された要因パターンの発生時点を付加する機能部である。QCD指標時系列情報付加部108は、QCD指標テーブル127の各々のレコードにQCD指標の算出の元となったQCD計測値に関する事象の発生時点を付加する機能部である。遅延相関算出部109は、時系列要因メトリクステーブル132と、時系列QCD指標テーブル133との情報を受け付けて、遅延相関リスト134を生成する機能部である。遅延相関分析による重要要因パターン分析部110は、実施形態1の重要要因パターン分析部106の機能に加え、遅延相関リスト134、要因パターン・要因メトリクス定義テーブル129、QCD計測値・QCD指標定義テーブル130の情報を受け付けて、遅延相関分析による重要要因パターン情報テーブル135を生成する機能部である。

【0110】

また、情報記憶部120に記憶されたデータに関しては、相関リスト128、重要要因

10

20

30

40

50

パターン情報テーブル 131 がなくなり、時系列要因メトリクステーブル 132、時系列 QCD 指標テーブル 133、遅延相関リスト 134、遅延相関分析による重要要因パターン情報テーブル 135 が付け加わっている。

【0111】

時系列要因メトリクステーブル 132 は、要因メトリクステーブル 125 に時点のフィールドが付け加わったテーブルである。時系列 QCD 指標テーブル 133 は、QCD 指標テーブル 127 に時点のフィールドが付け加わったテーブルである。遅延相関リスト 134 は、要因メトリクスと QCD 指標との遅延相関分析に関する係数を格納するリストである。

【0112】

次に、図 21 を用いて実施形態 2 に係るソースコード解析装置 10 で実行される処理の概要とデータフローについて説明する。

ソースコード解析装置 10 に、チケット 121、ソースコード変更履歴 122、ソースコード 123 が入力されることは、実施形態 1 と同様である。また、要因パターン取得部 101、要因メトリクス算出部 102、QCD 計測値取得部 103、QCD 指標算出部 104 の処理も実施形態 1 と同様である。

【0113】

本実施形態では、ソースコード解析装置 10 は、実施形態 1 の場合に加え、要因メトリクス時系列情報付加部 107 が、チケット 121、ソースコード変更履歴 122、ソースコード 123 と、要因メトリクステーブル 125 を受付け、また、時系列要因メトリクステーブル 132 を生成する。QCD 指標時系列情報付加部 108 が、チケット 121、ソースコード変更履歴 122、ソースコード 123 と、QCD 指標テーブル 127 を受付け、時系列 QCD 指標テーブル 133 を生成する。さらに、実施形態 1 で示した相関算出部 105 に代わり、遅延相関算出部 109 が、時系列要因メトリクステーブル 132、時系列 QCD 指標テーブル 133 を受付け、遅延相関リスト 134 を生成する。そして、実施形態 1 で示した重要要因パターン分析部 106 に代わり、遅延相関分析による重要要因パターン分析部 110 は、遅延相関リスト 134 を受付け、遅延相関分析による重要要因パターン情報テーブル 135 を生成する。

【0114】

次に、図 22 ないし図 25 を用いて実施形態 2 のソースコード解析装置で使用されるデータ構造について説明する。

【0115】

時系列要因メトリクステーブル 132 は、プログラム要素ごとに、要因メトリクスが計量された要因パターンの発生時点と、要因パターン属性テーブル 124 に格納されている要因パターンの属性に基づいたソフトウェア評価のための計量値を要因メトリクスとして格納するテーブルであり、図 22 に示されるように、プログラム要素名 132a、時点 132b、要因メトリクス 132ic_i (i = 1, 2, ...) のフィールドからなる。

【0116】

プログラム要素名 132a、要因メトリクス 132ic_i は、それぞれ要因メトリクステーブル 125 のプログラム要素名 125a、要因メトリクス 125ic_i (i = 1, 2, ...) と同様である。

【0117】

時点 132b には、当該レコードの要因メトリクスが計量された要因パターンの発生した時点が格納され、例えば、プログラムの変更時、チケットの更新時の時間情報などである。時点 132b の情報は、例えば、計測の間隔に従い、年、月、週、日、時など、計測間隔に従った単位で記録される。また、時点 1802 の記録形態は、実際の日付、または時刻として記録してもよく、計測の回次の番号として記録してもよい。

【0118】

時系列 QCD 指標テーブル 133 は、プログラム要素ごとに、元となる QCD 計測値テーブル 126 に格納されている QCD 計測値が計測された事象の発生時点と、QCD 計測

10

20

30

40

50

値に基づいた指標を格納するテーブルであり、図23に示されるように、プログラム要素名133a、時点133b、QCD指標 i 133c $_i$ ($i = 1, 2, \dots$) のフィールドからなる。

【0119】

プログラム要素名133a、QCD指標 i 133c $_i$ は、それぞれQCD指標テーブル127は、図7に示されるように、プログラム要素名127a、QCD指標 i 127b $_i$ と同様である。

【0120】

時点133bは、当該レコードのQCD指標を算出する元のQCD計測値を計測された事象の発生時点が格納され、例えば、不良が発生した時点、チケットによりプログラム開発の課題が明らかになった時点、工程遅延が発生した時点などの時間情報である。時点133bは、例えば、計測の間隔に従い、年、月、週、日、時など、計測間隔に従った単位で記録される。時点133bの記録形態は、実際の日付もしくは時刻として記録してもよく、計測の回次の番号として記録してもよい。

10

【0121】

遅延相関リスト134は、要因メトリクスとQCD指標との遅延相関分析に関する係数を格納するリストであり、図24に示されるように、要因メトリクス名134a、QCD指標名134b、加算係数134c、差分係数134d、係数134e、遅延相関係数134fの各フィールドからなる。

20

【0122】

要因メトリクス名134a、QCD指標名134bは、それぞれ相関リスト128の要因メトリクス名128a、QCD指標名128bと同様である。加算係数134c、差分係数134d、遅延係数134e、遅延相関係数134fは、当該レコードの要因メトリクスとQCD指標のペアに対する既に説明した遅延相関分析における加算係数、差分係数、遅延係数、遅延相関係数の値が格納される。

【0123】

遅延相関分析による重要要因パターン情報テーブル135は、要因パターンの種別、QCD計測値の種別ごとに、遅延相関分析の結果により算出された重要要因パターンを評価するスコアとそのスコアに基づいた要因パターンの重要性の順位を格納するテーブルであり、図25に示されるように、要因パターン種別名135a、QCD計測値名135b、影響観点135c、スコア135d、影響期間135eの各フィールドからなる。

30

【0124】

要因パターン種別名135a、QCD計測値名135b、影響観点135cは、それぞれ重要要因パターン情報テーブル131の要因パターン種別名131a、QCD計測値名131b、影響観点131cと同様である。スコア135dは、意味的には、重要要因パターン情報テーブル131のスコア131dと同じであるが、本実施形態では、遅延相関分析の結果のデータに基づいて算出されたスコアの値がスコア135dに格納される。影響期間135eには、当該レコードに関連する要因パターン種別とQCD計測値のペアに対する遅延相関分析の係数から算出される要因メトリクス(説明変数)、QCD指標(目的変数)の影響期間が格納される。

40

【0125】

次に、図26ないし図30を用いて実施形態2に係るソースコード解析装置が行う処理について説明する。

【0126】

先ず、図26を用いて要因メトリクス時系列情報付加処理について説明する。

ソースコード解析装置10の要因メトリクス時系列情報付加部107は、全てのプログラム要素Eに対して(S801-S803)、プログラム要素Eに対し要因メトリクスが計量された要因パターンの発生時点を取得し、時系列要因メトリクステーブル132に格納する(S802)。

【0127】

50

要因メトリクスが計量された要因パターンの発生時点は、チケット 121、ソースコード変更履歴 122、ソースコード 123 の時間情報から取り出すことができる。また、要因パターン取得部 101 が要因パターンの情報を取得したときに、要因メトリクス時系列情報付加部 107 にその情報を受け渡してもよい。また、ユーザが外部よりキーボードで時点に関する情報を与えるか、要因メトリクスと時点の対応テーブルを与えるようにしてもよい。

【0128】

次に、図 27 を用いて QCD 指標時系列情報付加処理について説明する。

ソースコード解析装置 10 の QCD 指標時系列情報付加部 108 は、全てのプログラム要素 E に対して (S901 - S903)、プログラム要素 E に対し QCD 指標の算出の元となった QCD 計測値に関する事象の発生時点を取得し、時系列 QCD 指標テーブル 133 に格納する (S902)。

【0129】

QCD 指標の算出の元となった QCD 計測値に関する事象の発生時点は、チケット 121、ソースコード変更履歴 122、ソースコード 123 の時間情報から取り出すことができる。また、QCD 指標算出部 104 が QCD 指標を算出したとき、QCD 指標時系列情報付加部 108 に、QCD 計測値に関する事象の発生時点の情報を受け渡してもよい。また、ユーザが外部よりキーボードで時点の情報を与えるか、QCD 指標と時点の対応テーブルを与えるようにしてもよい。

【0130】

次に、図 28 を用いて遅延相関算出処理について説明する。

ソースコード解析装置 10 の遅延相関算出部 109 は、全ての要因メトリクス E1 に対して (S1001 - S1005)、S1002 - S1004 の処理を行う。

S1001 - S1005 のループの中で、全ての QCD 指標 E2 に対して (S1002 - S1004)、遅延相関係数算出処理を行う (S1003)。遅延相関係数算出処理の詳細は、以下で説明する。

【0131】

次に、図 29 を用いて遅延相関係数算出処理の詳細を説明する。

これは、図 28 の S1003 に該当する処理である。

先ず、ソースコード解析装置 10 の遅延相関算出部 109 は、入力された対象となる QCD 指標および要因メトリクス値を正規化する (S1101)。例えば、(値の最大値 - 値の最小値) でそれぞれの値を除算することで値を正規化する。また、例えば、例えばボックスプロットを用いた外れ値除去をおこなう。

【0132】

全ての加算係数 E1 に対して (S1102 - S1108)、S1103 - S1107 の処理を行う。

S1102 - S1108 のループで、全ての差分係数 E2 に対して (S1103 - S1106)、S1104 - S1106 の処理を行う。

S1103 - S1106 のループで、全ての遅延係数 E3 に対して (S1104 - S1106)、入力された対象となる QCD 指標および要因メトリクス値に対して、既に説明した手法により、加算係数 E1、差分係数 E2、遅延係数 E3 における遅延相関係数を算出する (S1105)。

【0133】

全ての加算係数、差分係数、遅延係数の組み合わせのなかから、遅延相関係数が最大となる組み合わせを求め (S1109)、遅延相関リスト 134 に格納する。

【0134】

次に、図 30 を用いて遅延相関分析による重要パターン分析処理について説明する。

ソースコード解析装置 10 の図 26 に示す遅延相関分析による重要要因パターン分析部 110 は、全ての要因メトリクス・QCD 計測値の組 E1、E2 に対して (S1201 - S1205)、S1202 - S1204 を行う。

10

20

30

40

50

【0135】

S1201 - S1205のループの中で、要因メトリクスE1に対応する要因パターン種別名および重みを、要因パターン・要因メトリクス定義テーブル129より取得し、要因パターンのスコアを、遅延相関リスト134に格納された遅延相関係数とその重みに基づき、算出する(S1202)。例えば、要因パターンのスコアは、遅延相関係数と重みの積とする。

【0136】

次に、QCD指標E2に対応するQCD計測値名および重みを、QCD計測値・QCD指標定義テーブル130より取得し、QCD指標のスコアを、遅延相関リスト134に格納された遅延相関係数とその重みに基づき、算出する(S1203)。例えば、QCD指標のスコアのスコアは、遅延相関係数と重みの積とする。

10

【0137】

次に、遅延相関リスト134の加算係数と差分係数と遅延係数から、影響期間を算出する(S1204)。例えば、加算係数と遅延係数の和を影響期間とする。

【0138】

次に、全ての同じ要因パターン・QCD計測値の組Eに対して(S1206 - S1210)、S1207 - S1209を行う。

【0139】

S1206 - S1210のループの中で、要因パターンのスコアとQCD指標のスコアの和を取り、要因パターン・QCD計測値の組Eに対するスコアとし、同じ要因パターン・QCD計測値の組Eに対するスコアを合算し(S1207)、合算したスコアを遅延相関分析による重要要因パターン情報テーブル135に格納する。合算方法は、例えば、合計をとる。

20

【0140】

同じQCD計測値・要因パターンの組Eを有するS1260で算出された影響期間について、代表値を計算し、同じQCD計測値・要因パターンEの影響期間とし(S1208)、その影響期間を遅延相関分析による重要要因パターン情報テーブル135に格納する。例えば、S1260で算出された影響期間についての中央値を代表値とする。

【0141】

以上、本実施形態によれば、要因パターンを説明変数、QCD計測値を目的変数とした遅延相関分析を行うことにより、時間的順序を考慮したソフトウェア開発におけるアンチパターン、リファクタリングパターンによる影響の因果関係を定量的に把握することができ、ソフトウェアの品質向上、開発効率の向上を期待することができる。

30

【符号の説明】

【0142】

10...ソースコード解析装置、101...要因パターン取得部、102...要因メトリクス算出部、103...QCD計測値取得部、104...QCD指標算出部、105...相関算出部、106...重要要因パターン分析部、107...要因メトリクス時系列情報付加部、108...QCD指標時系列情報付加部、109...遅延相関算出部、110...遅延相関分析による重要要因パターン分析部、120...情報記憶部、

40

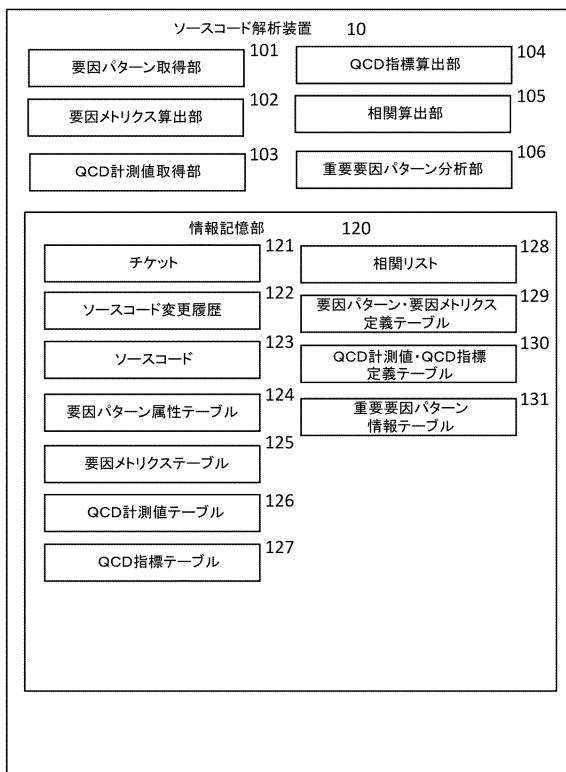
121...チケット、122...ソースコード変更履歴、123...ソースコード、124...要因パターン属性テーブル、125...要因メトリクステーブル、126...QCD計測値テーブル、127...QCD指標テーブル、128...相関リスト、129...要因パターン・要因メトリクス定義テーブル、130...QCD計測値・QCD指標定義テーブル、131...重要要因パターン情報テーブル、132...時系列要因メトリクステーブル、133...時系列QCD指標テーブル、134...遅延相関リスト、135...遅延相関分析による重要要因パターン情報テーブル

50

【図面】

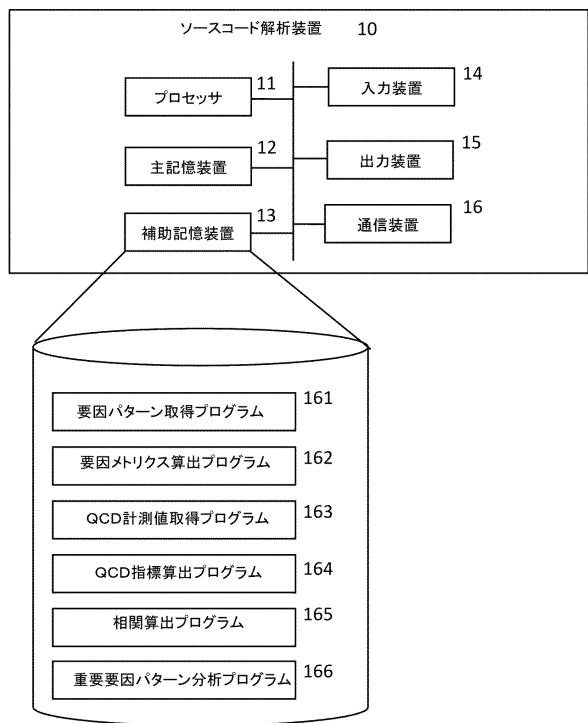
【図1】

図1



【図2】

図2



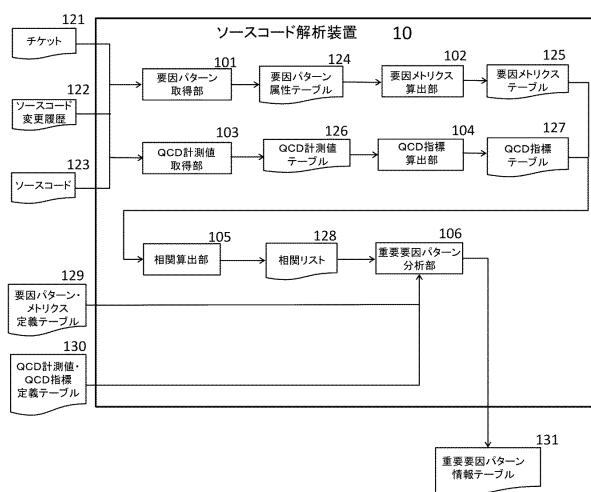
10

20

30

【図3】

図3



【図4】

図4



40

50

【図 5】

図5

| 要因メトリクステーブル 125 | | | |
|-----------------|-------------------|-------------------|-----|
| 125a | 125b ₁ | 125b ₂ | ... |
| プログラム要素名 | 要因メトリクス1 | 要因メトリクス2 | ... |
| プログラム要素A | 0.5 | 2 | |
| プログラム要素B | 0.3 | 2 | |
| プログラム要素C. | 0.8 | 1 | |
| プログラム要素D | 1 | 1 | |

【図 6】

図6

| QCD計測値テーブル 126 | | | | |
|----------------|------|---------|-------------------|-----|
| 126a | 126b | 126c | 126d ₁ | ... |
| QCD計測値種別 | 計測ID | プログラム要素 | 計測値1 | ... |
| チケット計測値1 | 1 | A | 10 | |
| チケット計測値1 | 1 | B | 10 | |
| チケット計測値1 | 2 | C | 5 | |
| 変更履歴 計測値1 | 1 | D | 1 | |

10

【図 7】

図7

| QCD指標テーブル 127 | | | |
|---------------|-------------------|-------------------|-----|
| 127a | 127b ₁ | 127b ₂ | ... |
| プログラム要素名 | QCD指標1 | QCD指標2 | ... |
| プログラム要素A | 10 | 2 | |
| プログラム要素B | 10 | 2 | |
| プログラム要素C. | 5 | 1 | |
| プログラム要素D | 10 | 1 | |

【図 8】

図8

| 相関リスト 128 | | |
|-----------|--------|------|
| 128a | 128b | 128c |
| 要因メトリクス名 | QCD指標名 | 相関係数 |
| 要因メトリクス1 | QCD指標1 | 0.8 |
| 要因メトリクス1 | QCD指標2 | 0.5 |
| ... | | |
| | | |
| | | |

20

【図 9】

図9

| 要因パターン・要因メトリクス定義テーブル 129 | | |
|--------------------------|----------|------|
| 129a | 129b | 129c |
| 要因パターン種別名 | 要因メトリクス名 | 重み |
| アンチパターン1 | 要因メトリクス1 | 0.8 |
| アンチパターン1 | 要因メトリクス2 | 0.2 |
| 改善パターン1 | 要因メトリクス3 | 1.0 |
| | | |
| | | |

【図 10】

図10

| QCD計測値・QCD指標定義テーブル 130 | | | |
|------------------------|--------|------|------|
| 130a | 130b | 130c | 130d |
| QCD計測値名 | QCD指標名 | 影響観点 | 重み |
| チケット計測値1 | QCD指標1 | Q | 0.8 |
| チケット計測値1 | QCD指標2 | Q | 0.2 |
| 変更履歴計測値2 | QCD指標3 | C | 1.0 |
| | | | |
| | | | |

40

50

【図 1 1】

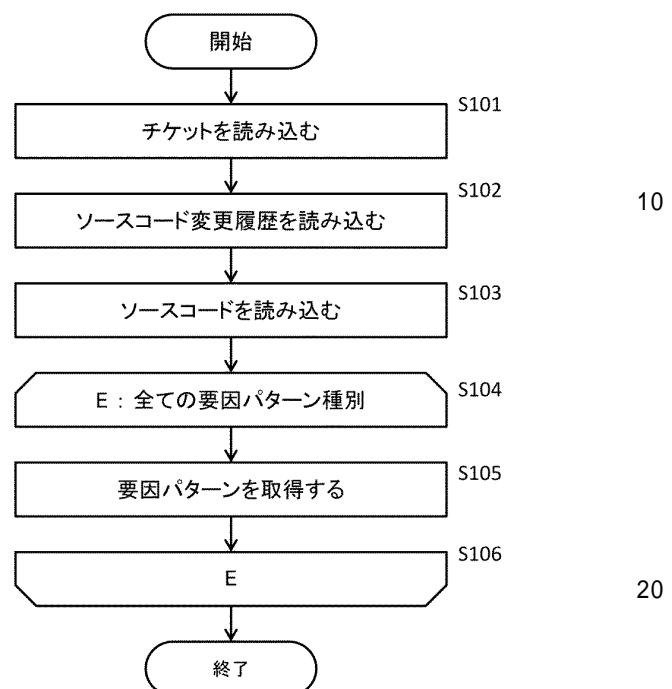
図 11

重要要因パターン情報テーブル 131

| 131a 要因パターン種別名 | 131b QCD計測値名 | 131c 影響観点 | 131d スコア | 131e 順位 |
|-------------------|-----------------|--------------|-------------|------------|
| アンチパターン1 | チケット計測値1 | Q | +100 | 1 |
| アンチパターン1 | 変更履歴計測値2 | C | +50 | 2 |
| | | | | |
| | | | | |
| | | | | |
| | | | | |

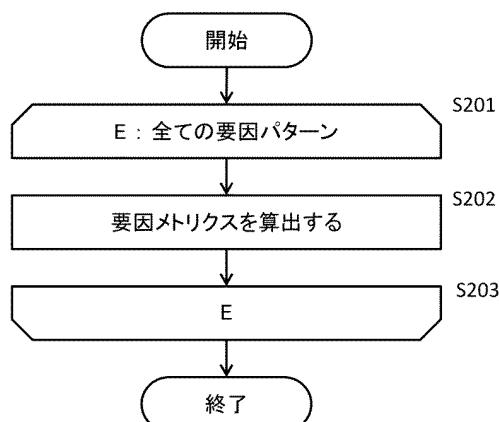
【図 1 2】

図 12



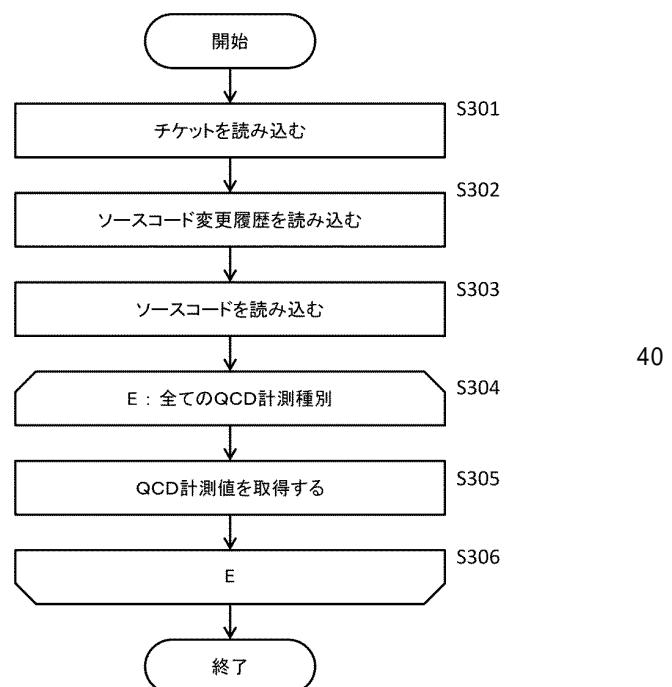
【図 1 3】

図 13



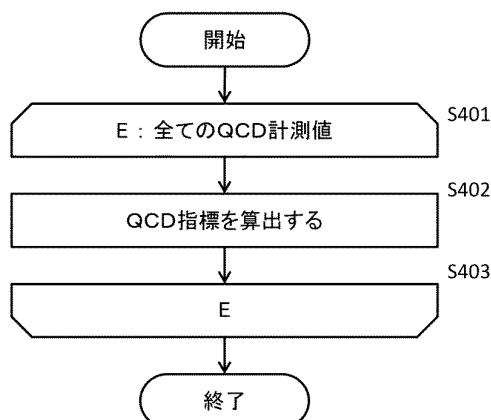
【図 1 4】

図 14



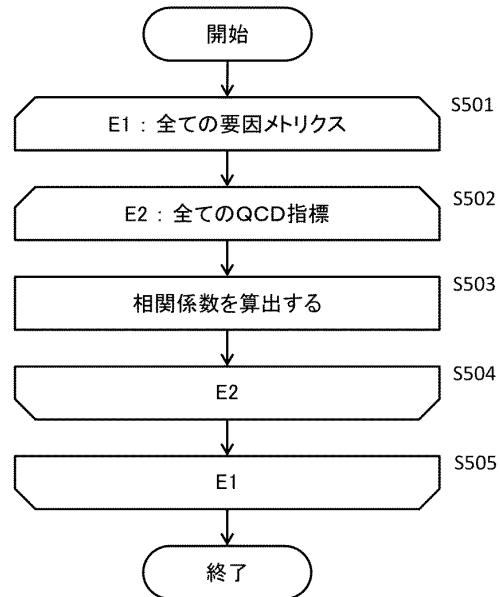
【図15】

図15



【図16】

図16

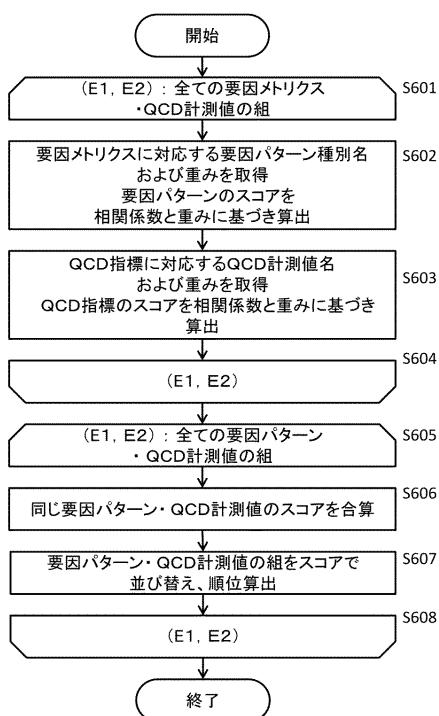


10

20

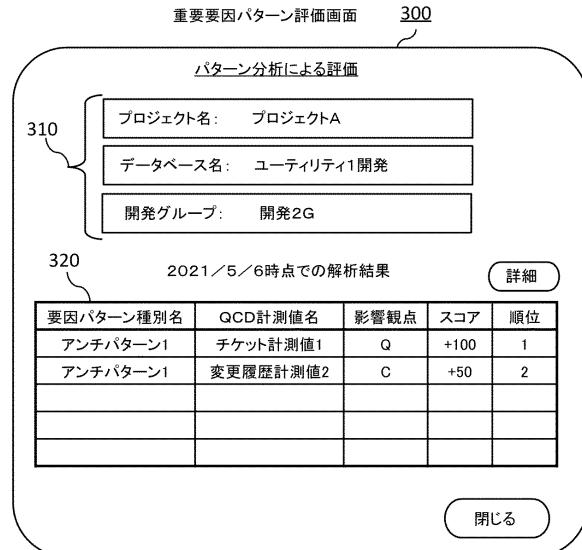
【図17】

図17



【図18】

図18

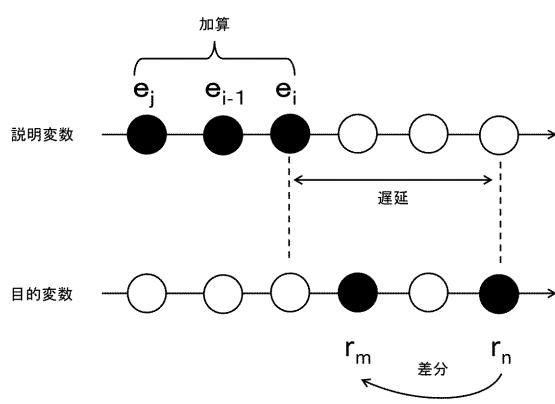


40

50

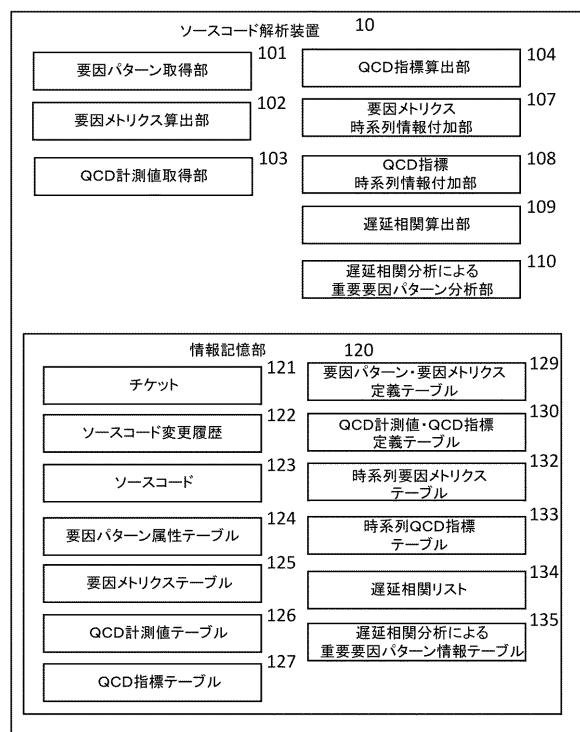
【図19】

図19



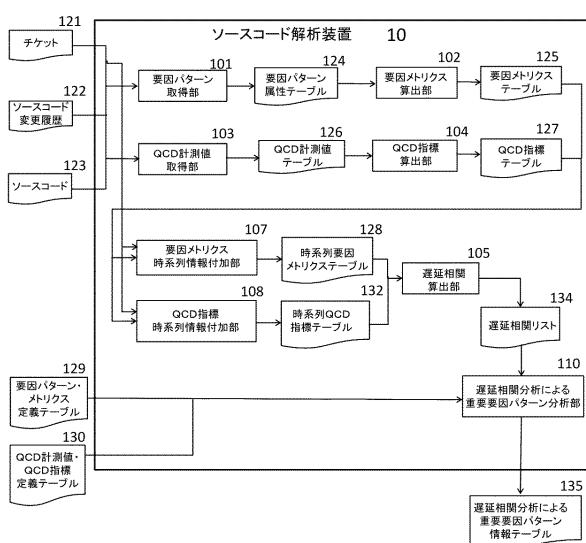
【図20】

図20



【図21】

図21



【図22】

図22



【図 2 3】

図23

時系列QCD指標テーブル 132

| 131a | 131b | 131c ₁ | 131c ₂ | ... |
|----------|--------|-------------------|-------------------|-----|
| プログラム要素名 | 時点 | QCD指標1 | QCD指標2 | ... |
| プログラム要素A | 2020/1 | 10 | 2 | |
| プログラム要素A | 2020/2 | 11 | 1 | |
| プログラム要素A | 2020/3 | 12 | 2 | |
| | | | | |
| | | | | |
| | | | | |

【図 2 4】

図24

遅延相関リスト 134

| 134a | 134b | 134c | 134d | 134e | 134f |
|----------|--------|------|------|------|--------|
| 要因メトリクス名 | QCD指標名 | 加算係数 | 差分係数 | 遅延係数 | 遅延相関係数 |
| 要因メトリクス1 | QCD指標1 | 3 | 2 | 2 | 0.8 |
| 要因メトリクス1 | QCD指標2 | 4 | 1 | 2 | 0.5 |
| ... | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |

10

【図 2 5】

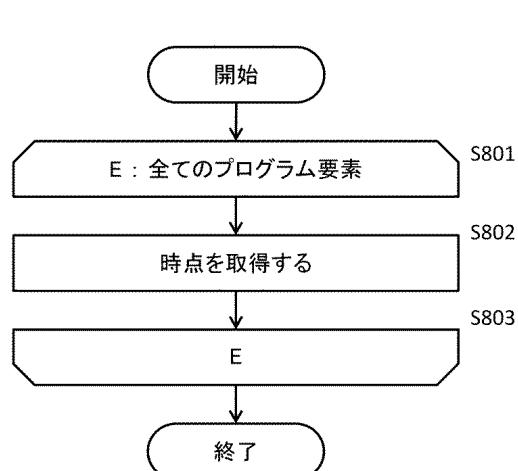
図25

遅延相関分析による重要要因パターン情報テーブル 135

| 135a | 135b | 135c | 135d | 135e |
|----------|----------|------|------|------|
| パターン種別名 | QCD計測値名 | 影響観点 | スコア | 影響期間 |
| アンチパターン1 | チケット計測値1 | Q | +100 | 3 |
| アンチパターン1 | 変更履歴計測値2 | C | +50 | 5 |
| | | | | |
| | | | | |
| | | | | |

【図 2 6】

図26



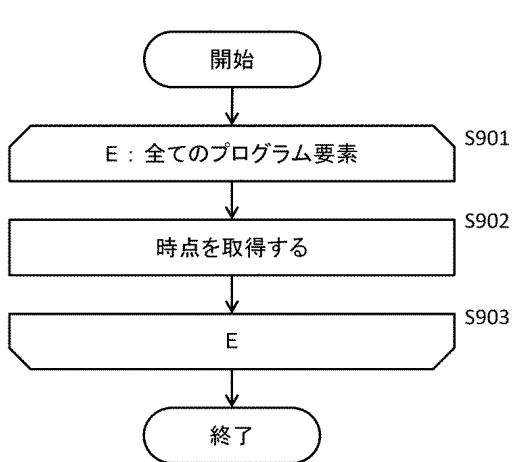
30

40

50

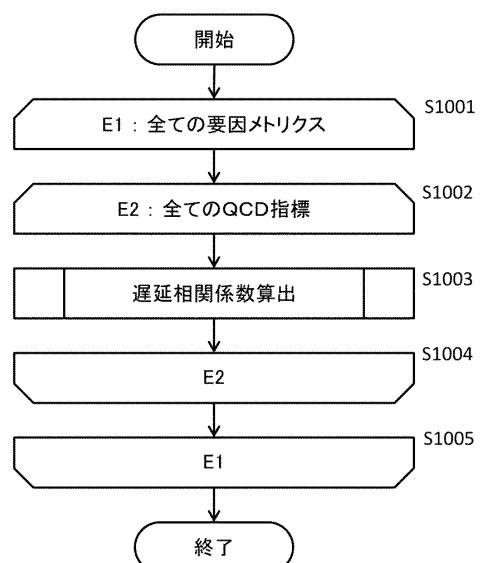
【図 27】

図27



【図 28】

図28

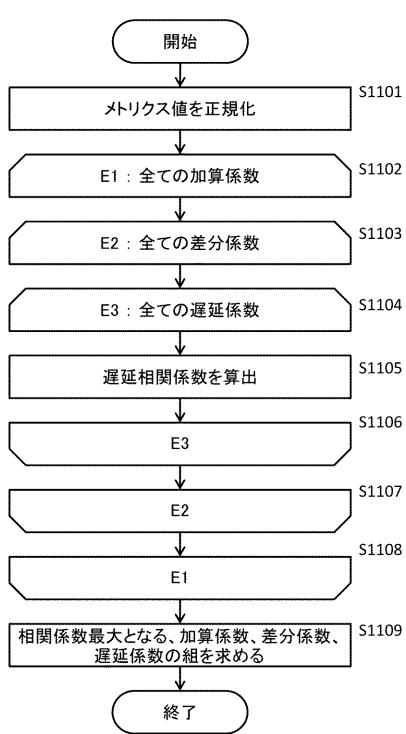


10

20

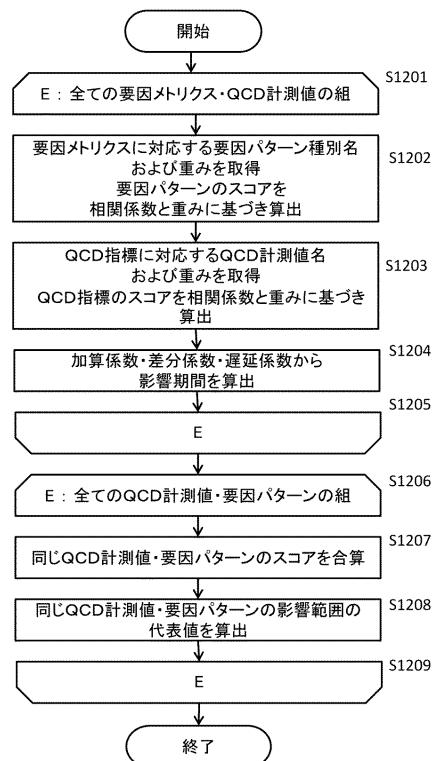
【図 29】

図29



【図 30】

図30



30

40

50

フロントページの続き

(56)参考文献 特開平07-160495(JP, A)

特開2014-222440(JP, A)

特開2003-050723(JP, A)

(58)調査した分野 (Int.Cl., DB名)

G 06 F 8 / 00 - 8 / 77