

ソフトウェア開発データ分析 モデリング言語の提案

角田雅照[†] 松村知子[†] 松本健一[†]

本稿では、ソフトウェア開発データ分析モデリング言語を提案する。GQM パラダイムや測定情報モデルでは、(1)ユーザ・ベンダ間でのデータ共有、(2)プロジェクト進行中のデータ分析、(3)分析結果に基づいた対応策の実施、(4)データ分析モデルの再利用、が考慮されておらず、十分なモデリングができないという問題がある。そこで上記を考慮して、測定情報モデルにユーザゴール、ベンダゴール、KGI、KPI、対応策、開始時条件、終了時条件などの概念を加えたモデリング言語を提案する。

A Modeling Language for Software Development Data Analysis

Masateru Tsunoda[†] Tomoko Matsumura[†] Ken-ichi Matsumoto[†]

We propose a new modeling language for software development data analysis. The GQM paradigm and the measurement information model is not sufficient for modeling software development data analysis because they do not consider the following issues: (1) data sharing between the user and the vendor, (2) analysis of projects in progress, (3) actions based on analysis results, (4) reuse of data analysis model. Considering the issues, we propose a modeling language based on the measurement information model, by adding user goal, vendor goal, KGI, KPI, actions, precondition, and post condition.

1. はじめに

近年、ソフトウェア開発の大規模化にともない、ソフトウェア開発プロジェクトの失敗時に、ユーザ、ベンダ双方が多大な損失を受ける危険性が高まっている。このため、ユーザ側でもプロジェクトの状況を把握したいという要望が高まっている。ユーザが進行中のプロジェクトの状況を正確に把握するためには、定量的データを計測、分析した結果をベンダと共有する必要がある。

ユーザ・ベンダ間でのデータ共有に先立って、データ分析の目的と方法を明確にし、合意(契約)する必要がある。GQM パラダイム¹や測定情報モデル²を用いてモデリングすることは可能であるが、以下の点が考慮されておらず、十分なモデリングができないという問題がある。

- (1) ユーザ・ベンダ間でのデータ共有
- (2) プロジェクト進行中のデータ分析
- (3) 分析結果に基づいた対応策の実施
- (4) データ分析モデルの再利用

そこで本稿では、上記を考慮したソフトウェア開発データ分析モデリング言語を提案する(図1)。

2. 提案するモデリング言語の概要

2.1 測定情報モデル

提案するモデリング言語は、測定情報モデル²をベースとする。測定情報モデルとは、データの計測から分析方法までを階層構造により表したモデルである。測定情報モデルには以下の要素が含まれる。

- 基本測定量: プロジェクト中に存在する属性を、定められた測定方法に基づいて定量化した数値。
- 導出測定量: 測定関数に基本測定量を与えることにより求められる数値。
- 管理指標: 分析モデルに導出測定量を与えることにより求められる数値。
- 解釈モデル: 管理指標の分析方法を示したモデル。

2.2 ユーザゴール, ベンダゴール

測定情報モデルには、データの計測から分析までのみがモデルに含まれており、データ分析の目的は含まれていない。そこで提案するモデリング言語では、測定情報モデルに GQM パラダイム¹のゴールの概念を加え、ゴールと管理指標を対応付ける。GQM パラダイムはデータ収集の目標設定からデータ収集のメトリクス決定までをモデル化したものである。GQM パラダイムにおけるゴールとは、計測の目標、計測対象、計測理由などを明確にした文である。さらに提案モデルでは、ユーザ・ベンダ間でデータを共有することを

[†] 奈良先端科学技術大学院大学情報科学研究科
Graduate School of Information Science, Nara Institute of Science and Technology

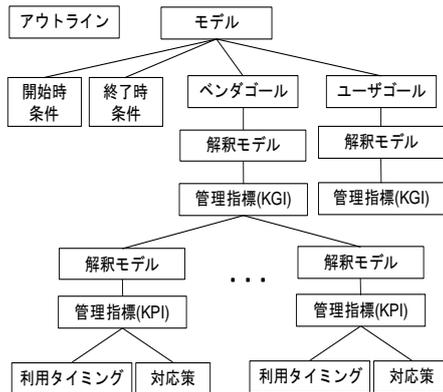


図1. ソフトウェア開発データ分析
モデリング言語の構成要素

考慮し、ゴールをユーザゴールとベンダゴールに分ける。

2.3 KGI, KPI

測定情報モデルでは、データ分析のゴールを達成したと判断する指標、ゴールの達成に向けてプロジェクトの進捗が順調であるかどうかを判断する指標が明確に定義されておらず、進行中プロジェクトのデータ分析方法をモデリングするには不十分である。そこで提案モデルでは、測定情報モデルに、ビジネスマネジメントの分野で用いられている KGI(Key Goal Indicator; 重要目標達成指標)、KPI(Key Performance Indicator; 重要業績評価指標)の概念を導入し、管理指標を以下のように KGI と KPI に分類する。

- KGI: ゴールを達成したか否かを判断するための定量的指標。
- KPI: プロジェクト進行中にプロセスを評価するための定量的指標。

KPI が目標値をクリアするようにプロジェクトを遂行することにより、KGI も目標値をクリアできるような関係となる。KGI はユーザゴール、ベンダゴールごとに(通常)1 つ定義され、KPI は KGI ごとに複数定義される。

2.4 対応策

測定情報モデルに対処策を加える。ゴールを達成するためには、分析結果に基づいて何らかの対応策をとる必要があるが、測定情報モデルには対応策が含まれていない。指標の分析結果を場合分けし、それぞれの場合における対応策を記述する。対応策は、KPI と KGI の組み合わせごとに記述する。例えば、KPI が「システムテスト欠陥密度」である場合、対応策は以下のように記述する。

- KPI が基準値よりも高かった場合
 - 単体、結合テストが適切かどうかを確認する。
- KPI が基準値よりも低かった場合
 - テストケースが適切かどうかを確認する。

2.5 開始時条件, 終了時条件

作成したデータ分析モデルを再利用するためには、そのモデルがどんな場合に適用できるのか、また、そのモデルの適用した分析結果に誤りがないかどうかを確認する必要がある。再利用時にこれらを確認できるように、以下のような開始時条件、終了時条件を測定情報モデルに加える。

- 開始時条件: 分析モデルを適用するために、プロジェクトがあらかじめ満たしておくべき条件。
 - 例: 要員のスキルが一定以上であるなど。
- 終了時条件: 分析モデル適用後に、解釈に誤りが含まれないことが保証される条件。
 - 例: 要件の頻繁な変更が発生しなかったなど。

2.6 アウトライン, 利用タイミング

データ計測、分析の動的な側面をモデリングできるように、アウトラインと利用タイミングを測定情報モデルに加える。アウトラインでは、分析モデルの適用場面、ユーザ/ベンダの要求、データ計測、分析手順が記述される。要求分析におけるユースケースと類似した内容を記述する。利用タイミングは、データの計測、データのユーザへの受け渡し、データ分析のタイミングを示した図である。UML のアクティビティ図を使って表記する。

3. まとめ

測定情報モデルをベースとした、ソフトウェア開発データ分析のためのモデリング言語を提案した。ワークショップでは、提案するモデリング言語の構成要素に過不足がないかどうかについて議論を行いたい。

謝辞 本研究の一部は、文部科学省「次世代 IT 基盤構築のための研究開発」の委託に基づいて行われた。

参考文献

- 1) Basili, V. and Weiss, D.: A Methodology for Collecting Valid Software Engineering Data, *IEEE Trans. On Software Eng.*, vol.10, No.3, pp.728-738, 1984.
- 2) International Organization for Standardization: ISO/IEC 15939:2002, Software Engineering—Software Measurement Process, International Organization for Standardization, Geneva, Switzerland, 2002.