Habilitationen

Andreas Henrich: Retrieval-Dienste für Software-Entwicklungsumgebungen

Anschrift: Universität – Gesamthochschule Siegen, Fachbereich Elektrotechnik und Informatik, Fachgruppe Praktische Informatik, D-57068 Siegen, Email: henrich@informatik.uni-siegen.de

1. Gutachter: Prof. Dr. Udo Kelter
2. Gutachter: Prof. Dr. Ralf Hartmut Güting
Tag der mündlichen Prüfung: 29. April 1997

Zusammenfassung: Software-Entwicklungsumgebungen (SEU) bestehen i.a. aus einer Reihe von Werkzeugen, mit denen Dokumente unterschiedlichster Typen bearbeitet werden können. Die Dokumente reichen von Analyse-Diagrammen über Design-Architekturen bis hin zu Programmtexten oder Testberichten. Sie werden im Rahmen der Software-Entwicklung nicht nur erstellt und editiert, sondern auch analysiert oder transformiert.

Um die Integration der Werkzeuge zu unterstützen, werden die Dokumente in modernen Umgebungen in einer gemeinsamen Datenhaltungskomponente — dem sogenannten *Repository* — abgelegt. Der Trend geht dabei hin zu einer feingranularen Modellierung der Daten im Repository. Damit ist eine gute Basis für werkzeugübergreifende Analysen, Qualitätssicherungsmaßnahmen oder Aufwandsschätzungen gegeben. Ferner wird z.B. auch die übergreifende, werkzeugunabhängige Suche nach wiederverwendbaren Komponenten möglich.

Das gesamte Potential eines solchen Repository zur Qualitätssicherung, Projektanalyse oder Redundanzvermeidung kann dabei nur ausgeschöpft werden, wenn auf dem Repository Retrieval-Dienste zur Verfügung stehen, die verschiedene Zugriffsparadigmen integrieren. Dies sind (1) das Navigieren durch den Objektgraphen, (2) der mengenorientierte Zugriff auf die Daten analog zu SQL/OQL und (3) die dokumentenorientierte Suche im Sinne des Information-Retrieval. Daneben ist offensichtlich eine Unterstützung dieser Zugriffsmöglichkeiten durch effiziente Zugriffsstrukturen erforderlich, um die für die praktische Anwendung erforderlichen Antwortzeiten zu gewährleisten.

In der vorliegenden Arbeit werden daher speziell auf die Erfordernisse des dargestellten Anwendungsbereichs zugeschnittene Retrieval-Dienste beschrieben, die die genannten Zugriffsparadigmen kombinieren. Ausgangspunkt der Überlegungen ist PCTE (Portable Common Tool Environment), der ISO- und ECMA-Standard zu Integrationsrahmen für SEU. PCTE spezifiziert u.a. ein operational objektorientiertes Objekt-Management-System zur Datenintegration. Aspekte des mengenorientierten Zugriffs oder der Dokumentensuche werden dabei aber nicht berücksichtigt.

Die Retrieval-Dienste werden in drei Hauptteilen betrachtet. Zunächst beschreiben wir eine an OQL angelehnte *Anfragesprache P-OQL* (PCTE Object Query Language), die auf dem Datenmodell von PCTE basiert und zusätzlich Dokumenten-Retrieval-Funktionalitäten (Zeichenkettensuche und Ähnlichkeitssuche nach dem Vektorraummodell) integriert. Zu dieser Anfragesprache wird auch eine formale Semantik angegeben.

Zur Implementierung von P-OQL schlagen wir eine mehrdimensionale *Zugriffsstruktur* vor, die Objekte gleichzeitig sowohl nach mehreren Standardattributen als auch nach Beschreibungsvektoren im Sinne des Vektorraummodells indexieren kann. Dadurch können kombinierte Anfragen, die eine Ähnlichkeitssuche mit Bedingungen über Standardattributen kombinieren, sehr effizient bearbeitet werden. In diesem Teil der Arbeit stellen wir auch einen Ansatz zur Optimierung von Anfragen vor, der spezielle Eigenschaften der Beziehungen in PCTE nutzt, um die zur Bearbeitung einer Anfrage am besten geeignete Indexstruktur auszuwählen.

Um die praktische Anwendbarkeit der Retrieval-Dienste darzustellen, werden schließlich vier *Beispielanwendungen* betrachtet: Ein Analyse-Werkzeug, ein Report-Generator, ein Werkzeug zur Schätzung des Entwicklungsaufwands von Software sowie ein Werkzeug zur Dokumentensuche.