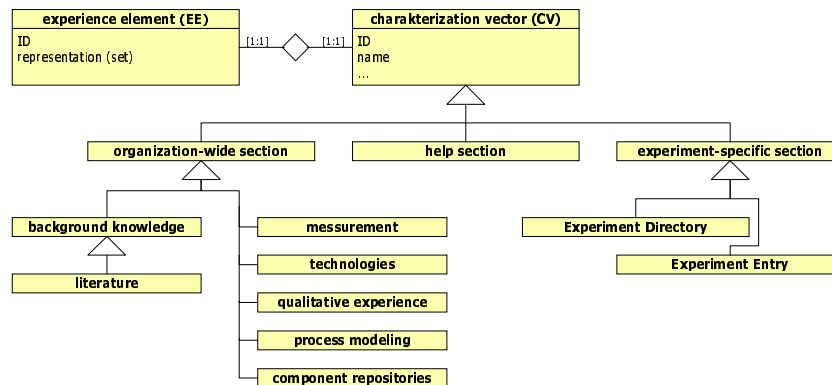


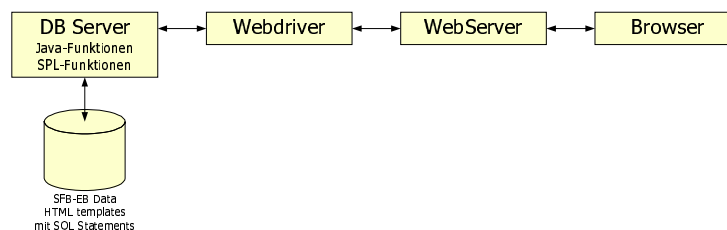
Unterstützung ähnlichkeitsbasierter Suche in der ORDB-gestützten SFB-501-Erfahrungsdatenbank

Der in den Teilprojekten A3 *Unterstützung des Softwareentwicklungsprozesses* und A1 *Software-Labor* des SFB's 501 *Entwicklung großer Systeme mit generischen Methoden* entwickelte und realisierte Ansatz einer ORDB-gestützten Erfahrungsdatenbank (EDB) unterstützt das systematische Ablegen und Wiederauffinden von während Softwareentwicklungsprozessen gewonnenen Erfahrungen (sog. Software-Artefakte). Damit stellt die EDB einen aktuellen Prototyp eines Reuse-Repositories dar.

Die eventuell in mehreren Repräsentationen vorliegenden Erfahrungsdaten werden als binäre Objekte in der objekt-relationalen Datenbank eingelagert und mit einem sog. Charakterisierungsvektor (CV) versehen, der den zugehörigen Erfahrungsdateneintrag näher beschreibt. Entsprechend der unterschiedlichen Erfahrungselementtypen (je nach Anwendungsdomäne) existieren auch CV-Typen, die in einer Generalisierungshierarchie organisiert sind.



Die Definition der CVs geschieht dabei mit Hilfe dynamisch generierter Eingabemasken oder durch beschreibende XML-Dateien. Die so erzeugten CVs dienen als Grundlage für eine ähnlichkeitsbasierte Suche nach potentiell wiederverwendbaren Software-Artefakten. Entsprechende rollen- und prozeßphasen-spezifische Funktionen werden den Benutzern über eine Web-Schnittstelle angeboten. Dazu greift der Web-Server über ein CGI-Programm (webdriver) auf den DB-Server (Informix) zu. Dort liegen entsprechende Web-Seiten, die zunächst durch Ausführung von benutzerdefinierten Java- und SPL-Funktionen vervollständigt und schließlich an den Browser geschickt werden.



Insbesondere in den frühen Phasen eines Softwareentwicklungsprozesses spielt die ähnlichkeitsbasierte Suche eine entscheidende Rolle, da sie das Auffinden und damit die Wiederverwendung von bereits zu früheren Zeitpunkten abgelegte, aber trotzdem im Hinblick auf das aktuelle Projekt hilfreiche Erfahrungen ermöglicht. Wie bereits angesprochen werden Erfahrungen dabei über beschreibende CVs aufgesucht. Durch Angabe eines Knotens in deren Generalisierungshierarchie bestimmt der Benutzer den Suchbereich in der Datenbank. Dies kann natürlich nur im Rahmen seiner jeweiligen Benutzerrechte geschehen.

Nach der Angabe des Suchbereichs kann nun eine Suchmaske generiert werden. Dazu müssen alle Attribute mit Hilfe des DB-Katalogs bestimmt werden, die im Suchbereich definiert sind und die der Benutzer lesen darf (Berücksichtigung von gesetzten Benutzerrechten). Danach erfolgt für jedes Attribut eine Bestimmung des Datentyps, die Generierung eines entsprechenden Eingabefeldes und eines Schaltfeldes zur Bearbeitung der Suchparameter für dieses Attribut. Nachdem Eingabefelder für die Mindestähnlichkeit der Ergebnisselemente und deren Sortierung erzeugt wurden, kann die Suchmaske angezeigt werden.

In der Abarbeitung einer ähnlichkeitsbasierten Suchanfrage wird bei jedem Vergleich einer in der Datenbank liegenden CV-Instanz mit der vom Benutzer angegebenen Vergleichsinstanz folgendermaßen vorgegangen. Für jedes Attribut des betrachteten CV-Typs existiert eine vom jeweiligen Datentyp abhängige Ähnlichkeitsfunktion (realisiert als benutzerdefinierte Java-Funktion, Ausführung im DB-Server), die die Ähnlichkeit des Attributwertes mit dem zugehörigen Vergleichswert berechnet. Die auf diese Weise errechneten Ähnlichkeitswerte der einzelnen Attribute werden zu einer Gesamtähnlichkeit des betrachteten CVs zusammengefaßt (jedes Attribut kann zusätzlich gewichtet werden). Bei diesem Vorgehen benötigt jede Ähnlichkeitsfunktion spezifische Parameter; hierfür werden Vorgabewerte eingesetzt, jeder Benutzer kann diese jedoch mit persönlichen Werten überschreiben.

Um diese Funktionalität zu realisieren, wird für jede Suchanfrage dynamisch ein SQL-Konstrukt generiert, das durch Aufruf und Verarbeitung der oben beschriebenen Ähnlichkeitsfunktionen die sich qualifizierenden Charakterisierungsvektoren und deren Ähnlichkeiten bestimmt. Schließlich werden die Suchergebnisse mit den jeweiligen Ähnlichkeiten präsentiert, dabei kann jedes Erfahrungselement genauer angezeigt werden oder auf der Ergebnismenge weitergesucht werden.

In diesem Beitrag wurde insbesondere aufgezeigt, wie die beschriebene ähnlichkeitsbasierte Suche mit den Erweiterungsmechanismen eines objekt-relationalen Datenbanksystems realisiert und im Rahmen einer Web-Schnittstelle zur Verfügung gestellt werden kann.

Michael Peter Haustein