

Diplomprüfung

Gedächtnisprotokoll
Praktische Informatik

Prüfer: Prof. Jarke

Fächer: EDB, IDB, Betriebssysteme

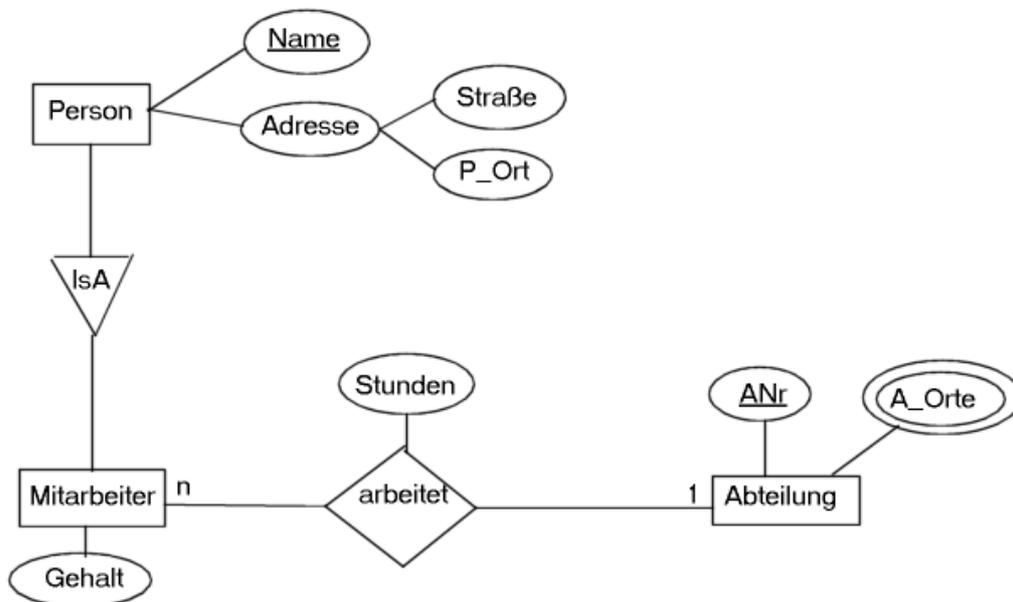
Datum: 24.04.2007

Name: Laszlo Bardos

Note: 1,3

Einstieg mit EDB:

- Datenbank Entwurf + Lebenszyklus erklären
- Ich soll jetzt konzeptuellen Entwurf mit ER Modell realisieren. Prof. Jarke gibt keine Vorgaben darüber was ich Modellieren soll, aber bittet möglichst alle „Features“ des ER Modells einzubauen. Ich modelliere folgendes:



Erkläre alle Komponenten. Prof. Jarke ist zufrieden und fragt wie das nun in ein RDB umgesetzt werden kann. Person als Relation mit Name, Straße und P_Ort als Attribute. IsA Beziehung als neue Relation Mitarbeiter, die Name als Schlüssel bekommt. Prof Jarke fragt nach den anderen Möglichkeiten wie eine IsA Beziehung modelliert werden könnte. In Mitarbeiter baue ich noch die arbeitet Beziehung ein. Die Abteilung Relation besteht nur aus ANr. Das A_Orte Attribut wird als eigenständige Relation realisiert weil mehrwertig. Die Relationen die ich erhalten habe waren also:

Person(Name, Straße, P_ort)

Mitarbeiter(Name, Gehalt, ANr, Stunden)

Abteilung(ANr)

A_Orte(Anr, A_Orte)

Ich hatte einen kleinen Hänger bei der Begründung warum ich arbeitet in die Relation von Mitarbeiter reingezogen habe, aber Prof. Jarke bugsierte mich dann ich die richtige Richtung.

- Jetzt sollte ich eine SQL Anfrage machen die den Namen aller Mitarbeiter die in Aachen wohnen zurückgibt. Habe zuerst: SELECT Person.Name FROM Person,

Mitarbeiter, Abteilung WHERE Person.Name='Aachen' AND

Person.Name=Mitarbeiter.Name AND Mitarbeiter.ANr=Abteilung.ANr

- Prof. Jarke merkt an, dass die Lösung zwar richtig ist, aber optimaler sein könnte. Irgendwann fällt mir auf, dass ich den Join von Abteilung gar nicht brauche und schreibe dann: SELECT Person.Name FROM Person, Mitarbeiter WHERE Person.Name='Aachen' AND Person.Name=Mitarbeiter.Name
- Jetzt sollte ich diese Anfrage in der relationalen Algebra angeben.
- Prof. Jarke fragt jetzt nach verteilten Datenbanken. Fragmentation und Allocation und dem Trade-Off von Redundanz und Transaktionskosten erklärt. (Danke Sascha! ☺)

Weiter ging es mit IDB:

- Prof. Jarke frage wie ein Join denn implementiert werden kann. Ich erzählte von Loop, Sort-Merge, Index und Hash Join. Erklärte dabei immer kurz das Prinzip.
- Jetzt ging es in Richtung Semi Join und ich erzählte von gutartigen Ausdrücken, zykelfreien Quantgraphen und, dass Semi Join nur eine Teilmenge der ursprünglichen Relation zurückgibt. Prof. Jarke fragt nach „Wachstum“ von mehreren Joins gegenüber mehreren Semi Joins. Ich antworte exponentiell gegenüber linear. Es schien die Richtige Antwort gewesen zu sein.

Jetzt Wechsel zu BS:

- Speicherhierarchie. Ich wäre sehr erschrocken gewesen, hätte ich nicht von jemand der am selben Tag vor mir dran war erfahren, dass er auch nach Speicherhierarchie gefragt wurde. So hab ich alle Zeiten parat gehabt und erwähne den Faktor 100.000 Sprung von Hauptspeicher zu Festplatte.
- Paging. Ich erzähle von virtuellem Speicher, Frames, Pages und Demand Paging Strategien. Prof. Jarke fragt nicht nach, sondern möchte wissen wie jetzt genau ein Zugriff auf eine Seite erfolgt. Ich stochere links und rechts im Nebel bis wir auf Page Tables und Page Offsets, worüber ich nicht viel Wissen habe. Ich werde gefragt wie es aussieht wenn die Page im Speicher ist und dann wie es aussieht wenn es von der Platte geladen werden muss.
- Jetzt werde ich zu meiner großen Überraschung zu Disc Scheduling gefragt, wie da optimiert vorgegangen werden kann. Mir fällt aus irgendeinem heise Artikel Command Queueing ein und ich erkläre das. War wohl die richtige Antwort. Jetzt noch gefragt wie sequentielles Lesen und auch Schreiben verbessert werden kann wenn man mehrere Platten hat. Habe diese Sachen auch nicht gelernt, aber erzähle alles was ich noch aus den Verlässlich Verteilten System Vorlesungen über Striping wusste. Nur durch die VVS Vorlesung wusste ich wofür die Abkürzung RAID überhaupt steht. Werde noch gefragt wie man innerhalb eine Platte optimal vorgehen kann. Zylinder gleichzeitig lesen etc..

Wir kommen jetzt wieder zurück zu IDB:

- Warum kann es problematisch sein, wenn das Betriebssystem Pages auf die Platte swappt wenn ein DBS darauf läuft. Diese Frage hatte ich in einem anderen Protokoll schon mal gelesen gehabt, und antworte, dass es Probleme mit UNDO und REDO geben kann. Das DBS hat keine Kontrolle wann Daten in den Sekundärspeicher gelangen. Muss das ein bisschen genauer erklären. Reden noch über Before-Images und dann ist die Prüfung auch schon vorbei.

Fazit:

Die Prüfung verlief in sehr angenehmer Atmosphäre und habe keine fiesen Fragen gestellt bekommen. Wenn ich mal nicht weiter kam, gab Prof. Jarke kleine Hilfestellungen die meistens ausreichten, um auf die richtige Lösung zu kommen.