Prof. Dr. Stefan Brass
Institut für Informatik
Universität Halle

10. Juli 2003

Vorlesung Grundlagen des World Wide Web
— Klausur —

Name: ______________________________

Studiengang: ______________________________

Matrikelnummer: ______________________________

(Diese Daten werden zur Ausstellung des Leistungsnachweises benötigt.)

Aufgabe	Punkte	Maximale Punktzahl
1 (Internet, DNS, URIs, HTTP)		11
2 (HTTP Request)		4
3 (XML)		10
4 (SGML/HTML)		3
Summe		28

Anleitung

- Sie haben Zeit bis 13^{45} (d.h. ca. 1:25).
- Bitte schreiben Sie die Antworten in den vorgegebenen Platz. Falls der Platz nicht ausreicht, können Sie weitere Blätter oder die Rückseite verwenden, aber bitte markieren Sie klar, daß es noch eine Fortsetzung gibt. Schreiben Sie Ihren Namen und die Nummer der Aufgabe auf alle zusätzlichen Blätter.
- Bitte fragen Sie, wenn Sie die Aufgabe nicht verstehen, oder die Aufgabe Ihnen mehrdeutig erscheint.
- Die Klausur hat 9 Seiten. Bitte prüfen Sie die Vollständigkeit.
- Bitte nehmen Sie keine Klausur mit und auch keine Notizen über die Aufgaben. Sie bekommen die Klausur zurück sobald sie korrigiert ist.
- Sie können alle schriftlichen Aufzeichnungen verwenden, die vor Beginn der Klausur entstanden sind. Dagegen sind Computer (Laptops etc.) leider nicht zulässig.
- Bitte vermeiden Sie jedes Verhalten, das als Täuschungsversuch mißverstanden werden könnte. Achten Sie auf einen ausreichenden Abstand zu Ihren Nachbarn und setzen Sie sich nicht direkt neben die Teilnehmer, mit denen Sie vorher immer die Hausaufgaben zusammen abgegeben haben. Schauen Sie nicht in Richtung der Klausuren anderer Teilnehmer und sprechen Sie nicht mit anderen Teilnehmern, bis alle Klausuren abgegeben sind. Tauschen Sie keineswegs irgendwelche Papiere aus. Rufen Sie notfalls den Dozenten zur Kontrolle.
- Bitte stellen Sie sicher, daß ich Ihre Handschrift lesen kann. Verwenden Sie keine rote Tinte. Sie dürfen mit Bleistift schreiben, allerdings können Sie dann später nicht beweisen, daß Sie etwas Bestimmtes geschrieben haben. In jedem Fall sollten Sie aber später nachfragen, wenn Sie die Korrekturen nicht verstehen.
- Meine Zeitschätzung ist: Aufgabe 1 je Teilaufgabe 2 Minuten, d.h. ca. 22 Minuten insgesamt. Für Aufgabe 2 ca. 10 Minuten. Für Aufgabe 3 ca. 30 Minuten. Für Aufgabe 4 ca. 10 Minuten. Damit würden noch 13 Minuten Reserve verbleiben. Falls Sie Schwierigkeiten mit einer Aufgabe haben, bearbeiten Sie zunächst die anderen.
- Viel Erfolg!

Aufgabe 1 (Internet, DNS, URIs, HTTP) 11 Punkte

Bitte kreuzen Sie jeweils die korrekte Antwort an. Pro Teilaufgabe sollte nur genau eine Antwort korrekt sein. Falls die die korrekte Antwort nicht wissen, sollten Sie wenigstens raten: Bei keinem Kreuzchen bekommen Sie ebenso null Punkte wie bei einem falschen Kreuzchen.

a) Was ist ein RFC?

- ☐ Remote Function Call
- ☐ Veröffentlichung über das Internet, z.B. HTTP Standard.
- ☐ Protokoll auf der Ebene 3 des OSI-Referenzmodells.
- ☐ Programm zum Listen von DNS Einträgen.

b) Wie große Dateien können Sie in ungefähr einer Sekunde über eine normale Modem-Verbindung übertragen? Gehen Sie davon aus, daß die Dateien bereits komprimiert sind. Kreuzen Sie die Antwort an, die dem richtigen Wert am nächsten kommt.

- ☐ 400 Byte
- ☐ 4 KB (KB = Kilobyte, d.h. 1024 Byte)
- ☐ 400 KB
- ☐ 40 MB (Megabyte)
- ☐ 4 GB (Gigabyte)

c) Was ist der Zweck des ARP-Protokolls bzw. des `arp` Kommandos?

- ☐ Einen Domainnamen in die IP-Nummer abzubilden.
- ☐ Zu testen, ob ein Rechner erreichbar ist.
- ☐ Einen SPAM-Filter zu konfigurieren.
- ☐ Die Ethernet-Adresse zu einer IP-Nummer herauszufinden.

d) Können Sie als normaler Benutzer unter UNIX einen Web-Server starten?

- ☐ Ja, aber nur mit einer Port-Nummer größer als 1024.
- ☐ Ja, mit einer beliebigen freien Port-Nummer.
- ☐ Nein, das können nur Administratoren (`root`).

e) Wer wählt den Port für einen Server?

- ☐ Das Server-Programm beim passiven Eröffnen der Verbindung.
- ☐ Das Client-Programm beim aktiven Eröffnen der Verbindung.
- ☐ Das Betriebssystem weist einen beliebigen freien Port zu.

f) Ich glaube, einmal auf einem Plakat folgende URL gesehen zu haben

`http://www.spd.ortsverein-giessen.de/`

Halten Sie diese URL für plausibel?

- ☐ Die Reihenfolge der Domain-Komponenten ist komisch.
- ☐ Der Bindestrich ist in URLs verboten.
- ☐ So hat die SPD die beste Kontrolle über die Subdomains (zentraler SPD Nameserver).

g) Die Informatik der Uni Halle möchte ihre Domain selbst verwalten. Wer muß darüber informiert werden und zustimmen? Sie können davon ausgehen, daß das Rechenzentrum die Domain der Universität Halle verwaltet.

- ☐ Niemand außerhalb der Informatik. Der Nameserver muß natürlich auf allen Informatik-Rechnern eingetragen werden.
- ☐ Das Universitäts-Rechenzentrum.
- ☐ Der DENIC.
- ☐ Die ICANN.
- ☐ Alle Root-Nameserver.

h) Was ist der Unterschied zwischen URL und URI?

- ☐ Keiner. Es sind sowohl theoretisch wie praktisch Synonyme.
- ☐ URI ist ein Tippfehler. Es gibt nur URLs.
- ☐ Jede URI ist auch eine URL, aber umgekehrt gibt es URLs, die keine URIs sind.
- ☐ URLs sind eine Teilmenge der URIs.

i) Was für Daten können mit dem HTTP-Protokoll übertragen werden?

- ☐ Nur HTML.
- ☐ Beliebige, auch binäre Daten.
- ☐ Nur in ISO Latin 1 codierte Textdaten.

j) Was ist die häufigste Methode in HTTP?

- ☐ GET.
- ☐ PUT.
- ☐ SELECT.
- ☐ RETRIEVE.
- ☐ CALL.
- ☐ DO.
- ☐ EXEC.

k) Wie lang ist eine IPv6 Adresse?

- ☐ 16 Bit
- ☐ 32 Bit
- ☐ 48 Bit
- ☐ 64 Bit
- ☐ 128 Bit
- ☐ 256 Bit

Aufgabe 2 (HTTP Request) 4 Punkte

Schreiben Sie einen HTTP/1.1 Request für die Seite

`http://ada.informatik.uni-halle.de:8080/~brass/`

Geben Sie dabei noch folgendes an:

- Sie wollen lieber `application/postscript` als `text/html`.
- Sie verstehen Deutsch und Englisch.
- Sie wollen die TCP-Verbindung nach Auslieferung der Seite nicht sofort schließen.

Aufgabe 3 (XML) 10 Punkte

Entwerfen Sie eine XML DTD für einen Straßenbahn- und Bus-Fahrplan und schreiben Sie ein kleines Beispiel-Dokument. Zur Vereinfachung können Sie ignorieren, daß

- Busse und Straßenbahnen am Wochenende anders fahren,
- es innerhalb einer Linie oft noch kleinere Varianten gibt,
- eine Linie normalerweise eine Hin- und Rückrichtung hat (wenn es keine Ringlinie ist).

Folgende Daten sollen also erfasst werden:

- Es ist eine Liste von Linien abzuspeichern, und für jede Linie eine eindeutige Nummer, sowie die Reihenfolge der Haltestellen zusammen mit der Fahrzeit vom Start bis zu der jeweiligen Haltestelle. In den Beispiel-Daten speichern Sie bitte die Linien 1 und 5 ab, die beide an der Haltestelle “Heide” starten. Für die Linie 1 geben Sie noch als zusätzliche Haltestellen “Markt” (Fahrzeit 12 Minuten) und “Frohe Zukunft” (24 Minuten) an. Für die Linie 5 “Markt” (12 Minuten) und “Riebeckplatz” (22 Minuten).
- Außerdem ist für jede Linie eine Liste von Uhrzeiten abzuspeichern, zu denen ein Bus/eine Straßenbahn von der Starthaltestelle losfährt (daraus kann man dann die Ankunft an jeder Haltestelle der Linie berechnen). Sie sollen aber für jede solche Fahrt abspeichern, ob ein Niederflur-Fahrzeug eingesetzt wird (erleichtert das Einsteigen mit Rollstuhl oder Kinderwagen). Für die Linie 1 geben Sie 19:13 und 19:33 als Abfahrtzeiten an, für die Linie 5 19:36 und 19:42.
- Sie brauchen nicht die syntaktische Richtigkeit der Uhrzeiten sicherzustellen, wenn Sie das nicht wollen.
- Sie sollen aber sicherstellen, daß man keine Start-Zeiten zu nicht existierenden Linien angeben kann.

Es ist Platz für die Lösung auf der nächsten Seite.

Platz zur Lösung von Aufgabe 3

Aufgabe 4 (SGML, HTML) 3 Punkte

Gegeben sei folgende stark vereinfachte SGML DTD für HTML:

```
<!ELEMENT HTML  O O (HEAD, BODY)>
<!ELEMENT HEAD  O O (TITLE & META?)>
<!ELEMENT TITLE - - (#PCDATA)>
<!ELEMENT META  - O EMPTY>
<!ATTLIST META
          http-equiv NAME  #IMPLIED
          content    CDATA #REQUIRED
<!ELEMENT BODY  O O ((#PCDATA|P|EM|UL|H1)*)>
<!ELEMENT H1    - - (#PCDATA)>
<!ELEMENT P     - O ((#PCDATA|EM|UL)*)>
<!ELEMENT EM    - - (#PCDATA)>
<!ELEMENT UL    - - (LI+)>
<!ELEMENT LI    - O ((#PCDATA|P|EM|UL)*)>
```

Folgendes HTML-Dokument enthält (mindestens) drei Fehler. Finden Sie diese Fehler. Nehmen Sie dabei an, daß die obige DTD die HTML 4.01 strict DTD ist, was natürlich nicht stimmt. Ich habe mich bemüht, nur drei Fehler in das Dokument einzubauen. Wenn Sie weitere Fehler finden, bekommen Sie Extrapunkte. Andererseits ziehe ich Ihnen für jede falsche Angabe, die über die drei geplanten Fehler hinaus geht, einen Punkt ab. Deswegen sollten Sie mehr als drei Fehler nur angeben, wenn Sie ganz sicher sind.

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN"
        "http://www.w3.org/TR/html4/strict.dtd">
<HTML>
        <META http-equiv="Content-Type" >
        <TITLE>My <EM>first</EM> HTML document!</TITLE>
    <BODY>
        <H1>Greeting
        <P>Hello, world!
            <UL    >
                <LI>Erster Punkt
                <EM>Was ist denn das?</EM>
                <LI>Letzter Punkt</LI>
            </UL>
    </BODY>
</HTML>
```

1. ______________________________

2. ______________________________

3. ______________________________