Prof. Dr. Stefan Brass, Dr. Henning Thielemann
Institut für Informatik
Martin-Luther-Universität Halle-Wittenberg

25. Juli 2008

Vorlesung „Grundlagen des World Wide Web“ — Klausur —

Name: ______________________

Studiengang: ______________________

Matrikelnummer: ______________________

(Diese Daten werden zur Ausstellung des Leistungsnachweises benötigt.)

Aufgabe	Punkte	Max. Punkte	Zeit
1 (Internet, DNS, HTTP)		10	15 min
2 (HTTP)		8	15 min
3 (XML)		10	30 min
4 (XHTML)		6	15 min
Summe		34	75 min

Hinweise:

- Bearbeitungsdauer: 90 Minuten
- Skript, Bücher, Notizen sind erlaubt.
  Notebooks, PDAs, Handys etc. dürfen nicht verwendet werden.
- Die Klausur hat 10 Seiten. Bitte prüfen Sie die Vollständigkeit.
- Verwenden Sie weder Rot- noch Bleistift zum Bearbeiten der Aufgaben.
- Bitte benutzen Sie den vorgegebenen Platz. Wenn Sie auf die Rückseite ausweichen müssen, markieren Sie klar, dass es eine Fortsetzung gibt.
- Tauschen Sie keinesfalls irgendwelche Dinge mit den Nachbarn aus. Notfalls rufen Sie eine Aufsichtsperson zur Kontrolle.
- Bei Aufgabe 1 zum Ankreuzen sollten Sie wenigstens raten, wenn Sie die richtige Lösung nicht wissen.
- Fragen Sie, wenn Ihnen eine Aufgabe nicht klar ist!

Aufgabe 1 (Internet, DNS, HTTP) 10 Punkte

Bitte kreuzen Sie jeweils die korrekte Antwort an. Es ist genau eine Antwort pro Teilaufgabe richtig. (Mehr als ein Kreuz pro Teilaufgabe wird mit 0 Punkten bewertet.)

a) Wie kann man eine Webseiten-Weiterleitung einstellen?

- ☐ mit der `Follow`-Option in `robots.txt`
- ☐ mit der `Redirect`-Option in `.htaccess`
- ☐ mit einem `Link`-Tag im Kopf der weiterzuleitenden HTML-Seite
- ☐ mit dem `Caption-Side`-Element in Style-Sheets

b) Vor etwa zehn Jahren gab es das Problem, dass auf `http://www.informatik.uni-halle.de/` abgelegte Binärdateien beim Herunterladen wie Textdateien behandelt wurden, d.h. dass Zeilenvorschübe (LF) durch MS-DOS/Windows-typische Zeilenenden (CR+LF) ersetzt wurden. Bei Binärdateien von anderen Servern gab es das Problem nicht. Woran könnte das gelegen haben?

- ☐ Falsche MIME-Typen in HTTP-Response-Header
- ☐ Falsche Browsereinstellung, wie spezielle MIME-Typen zu verarbeiten sind
- ☐ Das Problem hätte mit Cascaded-Style-Sheets Level 2 (CSS2) gelöst werden können, aber die gibt es erst seit 1998.

c) Angenommen, Sie rufen von 2 verschiedenen Browserfenstern gleichzeitig verschiedene URLs von dem selben Server per HTTP ab (verschiedene Sitzungen/Sessions auf demselben Client). Wie kann der entsprechende Server unterscheiden, welchem Browserfenster (welcher Sitzung) er welche Daten schicken soll?

- ☐ Der Browser schickt die Anfrage jedesmal an einen anderen Port auf dem Server.
- ☐ Der Browser schickt die Anfrage jedesmal von einem anderen Port auf dem Client.
- ☐ Die verschiedenen Sitzungen können eindeutig über jeweils unterschiedliche IP-Adressen des Clients unterschieden werden.
- ☐ Jeder Browser schickt eine eindeutige, zufällige ID mit jedem HTTP-Request mit, an der jede Sitzung eindeutig erkannt werden kann.

d) Welches Programm würden Sie verwenden, um die Hardware-Adresse (MAC) eines anderen Hosts im gleichen Subnetz zu bestimmen?

- ☐ `ping`
- ☐ `whois`
- ☐ `arp`
- ☐ `netstat`
- ☐ `nslookup`

e) Sie möchten mit Hilfe des DNS den zuständigen Mailserver zu einer E-Mail-Adresse herausfinden. Welchen DNS-Resource-Record müssten Sie dazu abrufen?

- ☐ A
- ☐ SOA
- ☐ CNAME
- ☐ HINFO
- ☐ MX
- ☐ NS

f) Zu welchem Host und Port stellt ein Browser (mit aktivierten Proxy-Einstellungen: „`www-cache.uni-halle.de:3128`“) eine Verbindung her, nachdem Sie die folgende URL eingegeben haben: „`http://www.apostroph.de/galerie.php`“?

- ☐ Host: `www.apostroph.de`, Port: `80`
- ☐ Host: `www.apostroph.de`, Port: `8080`
- ☐ Host: `www.apostroph.de`, Port: `3128`
- ☐ Host: `www-cache.uni-halle.de`, Port: `80`
- ☐ Host: `www-cache.uni-halle.de`, Port: `8080`
- ☐ Host: `www-cache.uni-halle.de`, Port: `3128`

g) Welches Protokoll verwendet der Browser, um obige Verbindung herzustellen?

- ☐ `UDP`
- ☐ `IPX`
- ☐ `TCP/IP`
- ☐ `FTP`
- ☐ `SMTP`

h) In einer Meldung auf einer Nachrichtenseite wurde einst über einen Hackerangriff berichtet, bei dem es gelungen war, von Microsoft unbemerkt unter `www.microsoft.com.is.hijacked.by.crackerjoe.org` gegen Microsoft hetzende Inhalte zu veröffentlichen. Wie muss der Hacker vorgegangen sein?

- ☐ Er ist in Rechner von Microsoft eingedrungen und hat diese manipuliert.
- ☐ Er ist in verschiedene Root-Nameserver eingedrungen und hat diese manipuliert.
- ☐ Es war kein Hack. Jemand hat sich lediglich in seiner privaten Domäne eine passende Unterdomäne angelegt.

i) Wie lang ist eine IPv6 Adresse?

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

j) Was ist der Zusammenhang zwischen URL, URI und URN?

- ☐ URL und URI sind theoretisch wie praktisch Synonyme. URN ist der Vorläufer von beiden und heute nicht mehr gebräuchlich.
- ☐ Die Menge der URIs ist die Vereinigung der Menge der URLs und der Menge der URNs.
- ☐ Die Menge der URLs ist die Vereinigung der Menge der URIs und der Menge der URNs.
- ☐ Die Menge der URNs ist Teilmenge der URLs und die wiederum Teilmenge der URIs.

Aufgabe 2 (HTTP) — 8 Punkte

a) Schreiben Sie einen HTTP-1.1-Request für die URL `http://www.gurgel.org/` über den Proxy-Server `www-cache.uni-halle.de` auf Port 3128. Geben Sie sowohl den Aufruf von `telnet`, als auch den HTTP-Request an.

- Sie akzeptieren sowohl HTML (`text/html`) als auch reinen Text (`text/plain`) als Ausgabeformat, ziehen jedoch HTML vor.
- Sie wollen eventuelle Ausgabetexte (Fehlermeldungen etc.) in Deutsch („`de`“).
- Die Verbindung soll für weitere Anfragen offen gehalten werden.
- Es sollen nur die ersten 256 Bytes übertragen werden.

. . . .

. . . .

. . . .

. . . .

. . . .

. . . .

b) Der Proxy-Server antwortet mit:

```
HTTP/1.0 503 Service Unavailable
Server: squid/2.6.STABLE17
Date: Thu, 17 Jul 2008 21:31:04 GMT
Content-Type: text/html
Content-Length: 1212
Expires: Thu, 17 Jul 2008 21:31:04 GMT
X-Squid-Error: ERR_DNS_FAIL 0
X-Cache: MISS from www-cache.uni-halle.de
X-Cache-Lookup: MISS from www-cache.uni-halle.de:3128
Via: 1.0 www-cache.uni-halle.de:3128 (squid/2.6.STABLE17)
Proxy-Connection: close

<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN"
 "http://www.w3.org/TR/html4/loose.dtd">
<HTML><HEAD><META HTTP-EQUIV="Content-Type"
 CONTENT="text/html; charset=iso-8859-1">
...
```

- Konnte das Dokument im Cache gefunden werden und woran erkennen Sie das?

  . . . .

- Welche Länge hat das Ergebnisdokument und warum?

  . . . .

Aufgabe 3 (XML) 10 Punkte

Stellen Sie sich vor, Sie wollen eine Datenbank für `lobbycontrol.de` über Interessenkonflikte bei Politikern in Form einer XML-Datei erstellen.

- Jeder Politiker hat einen Namen und gehört einer Partei an. Gehen Sie von einer kleinen Auswahl von Parteien aus.
- Eine Interessenverflechtung besteht aus einer oder mehreren politischen Aktivitäten und einer oder mehreren privatwirtschaftlichen Aktivitäten auf dem gleichen Gebiet. Manche Aktivitäten lassen sich mit einem Dokument aus dem WWW belegen.
- Jeder Politiker unterliegt Interessenverflechtungen auf einer beliebigen Anzahl von Gebieten.

Entwerfen Sie einen XML-DTD und ein dazugehöriges XML-Dokument, welches folgende Daten enthält: Rudolf Scharping (SPD) hat am ÖPP-Beschleunigungsgesetz mitgewirkt und ist heute Geschäftsführer der „Rudolf Scharping Strategie Beratung Kommunikation GmbH“ (`www.rsbk.de`) mit Schwerpunkt auf Beratung bei ÖPP. [1] Otto Schily (Mitbegründer der Grünen, seit 1989 SPD, suchen Sie sich eine Partei aus) hat als Bundesinnenminister Pässe mit biometrischen Merkmalen eingeführt und sitzt seit August 2006 in den Aufsichträten der beiden Biometrie-Unternehmen SAFE ID Solutions AG und Biometric Systems AG. Klaus-Heiner Lehne (CDU) ist Europaparlamentarier und hat sich als Koordinator der EVP-ED-Fraktion im Rechtsausschuss massiv für unbegrenzte Patentierbarkeit eingesetzt. Gleichzeitig arbeitete er für die Anwaltssozietät Taylor Wessing, welche sich unter anderem mit Patentstreitigkeiten befasst. Walter Riester (SPD) hat die nach ihm benannte Rente eingeführt und hält heute Vorträge für jeweils mehr als 7000 Euro Honorar bei Anbietern von Riesterrenten und deren Vertretern, wie dem AWD (`www.awd.de`), dem Bundesverband Deutscher Vermögensberater und vielen anderen.

Verwenden Sie in dem XML-Dokument lediglich ein Stichwort pro Aktivität. Es ist Platz für die Lösung auf der nächsten Seite. Geben Sie DTD und Beispieldokument mit den oben aufgeführten Daten an, sowie ggf. Bedingungen, die oben genannt sind, die Ihre DTD aber nicht sicherstellt.

[1] Öffentlich-Private Partnerschaften (ÖPP) sind eine Form der Teilprivatisierung, wie sie auch in Halle zum Beispiel in einem „Pilot“-Projekt bei ausgewählten Kindergärten und Schulen und seit kurzem auch bei Wohnungen der HWG angewendet wird. Dabei werden öffentliche Einrichtungen an private Betreiber verkauft und für einige Jahrzehnte zurückgemietet. Die Kommune hat durch den Verkaufserlös zunächst nominal weniger Schulden, dafür aber langfristige, meist höhere Verpflichtungen aus den Mietzahlungen.

Platz zur Lösung von Aufgabe 3

Aufgabe 4 (XHTML) 6 Punkte

Gegeben sei folgende stark vereinfachte DTD für XHTML:

```
<!DOCTYPE html [

<!ELEMENT html        (head, body)>
<!ATTLIST html        xmlns CDATA #IMPLIED>
<!ENTITY  % head.misc "(meta|link)*">
<!ENTITY  % block     "p|h1|ul">
<!ENTITY  % inline    "#PCDATA|em|img">
<!ELEMENT head        (%head.misc;, title, %head.misc;)>
<!ELEMENT title       (#PCDATA)>
<!ELEMENT meta        EMPTY>
<!ATTLIST meta        http-equiv NMTOKEN #IMPLIED
                      name       NMTOKEN #IMPLIED
                      content    CDATA   #REQUIRED>
<!ELEMENT link        EMPTY>
<!ATTLIST link        href CDATA #REQUIRED
                      rel (start|next|prev) #IMPLIED>
<!ELEMENT body        (%block;)+>
<!ELEMENT h1          (#PCDATA)>
<!ELEMENT p           (%inline;)*>
<!ELEMENT em          (%inline;)*>
<!ELEMENT ul          (li)+>
<!ELEMENT li          (%inline;|%block;)*>
<!ELEMENT img         EMPTY>
<!ATTLIST img         src     CDATA  #REQUIRED
                      alt     CDATA  #REQUIRED
                      height  CDATA  #IMPLIED
                      width   CDATA  #IMPLIED>
]>
```

Auf der nächsten Seite folgt ein XHTML-Dokument mit (mindestens) 6 Fehlern. Finden Sie diese Fehler! Nehmen Sie dabei an, dass die obige DTD die XHTML-1.1-DTD ist, was natürlich nicht stimmt. Es sollten nur 6 Fehler in dem Dokument enthalten sein. Wenn Sie weitere Fehler finden, bekommen Sie Extrapunkte. Andererseits wird Ihnen für jede falsche Angabe, die über die 6 geplanten Fehler hinausgeht, ein Punkt abgezogen. Deswegen sollten Sie mehr als 6 Fehler nur angeben, wenn Sie ganz sicher sind.


```
01 | <?xml version="1.0" encoding="ISO-8859-1"?>
02 | <!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.1//EN"
03 |          "http://www.w3.org/TR/xhtml11/DTD/xhtml11.dtd">
04 | <html xmlns="http://www.w3.org/1999/xhtml">
05 |   <head>
06 |     <meta name="author" content="Günhild Gründünk" />
07 |     <title>Das Schweigen der Lemmata</title>
08 |     <link rel=next href="autor.html">&Uuml;ber die Autorin</link>
09 |   </head>
10 | <body>
11 |   <h1>Die List</h1>
12 |   <P>
13 |   Ein Bild <img src="held.jpg?width=42&height=23"></img>
14 |   <ul><li>Ein Unterpunkt</li><li/></ul>
15 | </body>
16 | </html>
```

Bitte tragen Sie Ihre Lösung in die folgenden Felder ein. Nutzen Sie dabei die angegebenen Zeilennummern für die Beschreibung des Fehlerortes.

....

....

....

....

....

....