

Name and initials:	Grade
Student-number:	Course: Web Technology
Year of 1st enrollment:	Date: 8 April 2013

The exam consists of 20 closed questions (5 points in total) and two open questions (2 points each). You get one point for free. Please try to avoid corrections on these forms, use additional paper for draft answers.

1 Closed questions (5 points)

Each question consists of four statements, and each statement can be true or false (that is, for each question, the number of true statements can vary from 0 to 4). Indicate below which statements are true:

1. Which of the following statements about the history and organization of the Web are true?

- ☐ a. W3C is a commercial (for profit) organization.
- ☐ b. The 404 error code (“Page not found”) has slowed down the growth of the early Web significantly
- ☐ c. The Internet Protocol originates from the early nineties
- ☐ d. The notion of “hypertext”, text with links, existed before the Web was invented

2. Which of the following statements about IP protocols are true?

- ☐ a. IPv6 (Internet Protocol version 6) is the successor of IPv4.
- ☐ d. 201.123.109.1 is a valid IPv4 address.
- ☐ c. A relatively large portion of the IPv4 address space is reserved for addresses in the US.
- ☐ b. The IPv4 address space can encode 2^{32} addresses, while IPv6 can encode 2^{64} addresses.

3. Which of the following statements about Internet protocols are true?

- ☐ a. The UDP protocol is built on top of the TCP protocol.
- ☐ b. The TCP protocol is built on top of the IP protocol.
- ☐ c. The TCP protocol is more reliable than the IP protocol.
- ☐ d. The TCP protocol introduces the notion of “port”.

4. Which of the following statements about HTML elements are true?

- ☐ a. A child of the **body** element must be a block element.
- ☐ b. An inline element can have a block element as its child.
- ☐ c. Inline elements are so-called “empty” elements.
- ☐ d. The **span** element is an inline element.

5. Which of the following statements about the “cascading” process of CSS are true?
- ☐ a. User agent style definitions have a higher priority than author style definitions.
 - ☐ b. User style definitions always have a higher priority than author style definitions.
 - ☐ c. User style definitions have a higher priority than user agent style definitions.
 - ☐ d. “important” author style definitions have a higher priority than “important” user style definitions.
6. Which of the following statements about HTTP are true?
- ☐ a. HTTP request and response messages are usually binary encoded.
 - ☐ b. HTTP is a “stateless” network protocol.
 - ☐ c. HTTP is part of a “RESTful” network architecture.
 - ☐ d. The HTTP request method *must* be either **GET** or **HEAD**.
7. Which of the following statements about cookies are true?
- ☐ a. Cookies are used to adapt sites to the personal preferences of the user.
 - ☐ b. Cookies are used to remember action users performed previously
 - ☐ c. Cookies can be used to pass on private and sensitive information.
 - ☐ d. Cookies often contain viruses
8. Which of the following statements about information laws and the structure of the Web are true?
- ☐ a. The average shortest path between two arbitrary sites on the Web is relatively small, this is mainly because of a few big sites that connect many other sites
 - ☐ b. The Web is a small world network, and this is mainly because it contains many pages, but relatively few hyperlinks.
 - ☐ c. Heaps’ law predicts that if you need to index many documents, you will have quickly encountered all unique words in the collection.
 - ☐ d. Zipf’s law predicts that in natural text, the most frequent word will occur roughly two times more often as the second most frequent word.
9. Which of the following statements about search engines are true?
- ☐ a. A document is highly relevant if and only if it contains the term you search on.
 - ☐ b. **tf.idf** ranks a document higher if it contains a search term that is rare in the entire document collection.
 - ☐ c. Search engines have an index, so they do not need to crawl the Web at query time.
 - ☐ d. Precision and recall statistics are absolute and are not based on human interpretation of search results
10. Which of the following statements about PHP are true?
- ☐ a. The way the PHP interpreter is built makes PHP an inherently safe language.
 - ☐ b. A PHP processor leaves HTML markup in a **.php** file unchanged, but replaces PHP code fragments by the output they produce.
 - ☐ c. PHP is often used in combination with a relational database management system such as MySQL.
 - ☐ d. PHP has no “print” or “echo” statement as part of the language.

11. Which of the following statements about JavaScript are true?

- ☐ a. Using JavaScript host objects, you can close the browser window.
- ☐ b. JavaScript has no “print” or “echo” statements to produce textual output.
- ☐ c. With JavaScript and the DOM API, you can remove an `img` element from the tree structure of an HTML document.
- ☐ d. With JavaScript and the DOM API you can modify the style of a `p` element.

12. Which of the following statements about scripting are true?

- ☐ a. The Document Object Model (DOM) models a document as a tree data structure.
- ☐ b. The DOM API can be used in JavaScript, but not in other programming languages.
- ☐ c. An event handler is a function that is called in response of a specific event.
- ☐ d. A click on your mouse or pressing a key on the keyboard are good examples of such an event.

13. Which of the following statements about AJAX are true?

- ☐ a. With AJAX technology, you can replace parts of the page without refreshing the entire page.
- ☐ b. AJAX technology makes inaccessible sites accessible.
- ☐ c. Websites no longer need to use JavaScript if they start using use AJAX.
- ☐ d. Websites no longer need to use HTML if they start using AJAX.

14. Which of the following statements about Internet protocols and languages are true?

- ☐ a. According to the W3C HTML Recommendation, JavaScript is the only allowed language in a `script` element.
- ☐ b. According to the W3C HTML Recommendation, PHP is the only allowed language in a `<? . . . ?>` “processing instruction”
- ☐ c. `ftp://example.org/resource.txt` is a syntactically valid URL.
- ☐ d. RSS is a vocabulary that can be used to publish the latest updates of a website.

15. Which of the following statements about accessibility are true?

- ☐ a. A Web document that has been successfully validated by the W3C HTML en CSS validator also conforms to the W3C accessibility guidelines.
- ☐ b. Appropriate use of the `alt` attribute improves the accessibility of video, audio and image content.
- ☐ c. Dynamic Web content is never a Web site accessibility problem.
- ☐ d. “Keyboard shortcuts” are an important threat for Web accessibility.

16. An evaluation study uses cookies to measure the duration of the session of each participant. The hypothesis tested is: “Older users stay longer on our site than younger visitors.” Which of the following statements about this study are true?

- ☐ a. The age of the participants is an independent variable in this study.
- ☐ b. The session duration of the participants is a dependent variable in this study.
- ☐ c. De server log files contain sufficient information to carry out this research, if the cookies are properly logged as well.
- ☐ d. This research is an example of a qualitative study.

17. Which of the following statements about website evaluation are true?

- ☐ a. You can make a “low-fidelity mockup” with pen and paper.
- ☐ b. A/B testing tests a variant of one aspect of a page.
- ☐ c. Log files of websites are not subject to privacy laws.
- ☐ d. With online surveys, you have less control over your population.

18. Which of the following statements related to Web Science are correct?

- ☐ a. The laws of the physical world typically also apply in the virtual world.
- ☐ b. Legal issues on the Web are extra difficult due to lack of “location”.
- ☐ c. Web Science is Computer Science applied to the Web.
- ☐ d. The complex structure of the Web is caused by the complexity of Web protocols.

19. Net neutrality implies that:

- ☐ a. the Internet should be free of political interference.
- ☐ b. the Internet should be free of charge.
- ☐ c. people should have equal opportunities to access information on the Internet.
- ☐ d. the Internet should take a neutral position in political debates.

20. Which of the following statements about security are true?

- ☐ a. According to Kerckhoff’s principle one should keep the key secret, and make all algorithms public.
- ☐ b. Alice and Bob can use public-key encryption to keep their information secret, but not to verify that their information has not been changed by a man-in-the-middle attack.
- ☐ c. RSA public key encryption allows the use of the public key for encoding, and the private key for decoding, and *vice versa*
- ☐ d. Authentication based on an iris scan is a good example of “two-factor” authentication.

2 Open questions

Question 21 (2 points)

Study the HTML code on the next page (source 1)

- a) Identify (at least 5) errors against the HTML *syntax*, and indicate how the markup needs to be improved to make the code valid XHTML Strict.
- b) Identify (at least 5) cases where the content and style markup are syntactically valid, but are still logically or semantically incorrect. Again, indicate how these errors can be resolved.

Indicate clearly which answers belong to a) and which belong to b).

Answer:

Source 1

```
1 <!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
2   "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
3 <html xmlns="http://www.w3.org/1999/xhtml">
4
5 <head>
6   <style>
7     body {background-color: #FFFFFF;}
8     div { /* this is for the first level heading */
9         font-weight: bold;
10        font-size: 40pt;
11      }
12   </style>
13   <title>WT - Assignment 1</title>
14 </head>
15
16 <body style="width: 1024px; height: 768px;" bgcolor="silver">
17   <div style="font-size: 30pt; font-weight: bold;">
18     Web Technology: Assignment 1 <!-- this is the level first heading -->
19   </div>
20
21   <p style="text-align: center;">
22     The purpose of <strong>assignment 1</strong> is to test whether
23     each <span style="font-weight: bold;">group</span>
24     has grasped the key concepts of XHTML at a sufficient level.
25   <blockquote>
26     We expect that students who have actively participated in
27     making this assignment will not find the exam very difficult.
28   </blockquote>
29   </p>
30
31   
32
33   <p style="text-align: center;">
34     You are expected to use a variety of HTML elements and
35     attributes to get a grade > 5.0.
36   <p style="text-align: center;">
37     Please do not forget to test the CSS <code>font-family</code> property!
38   </p>
39
40 </body>
41 </html>
```

Question 22 (max 2 points)

Study the 7 lines from the server log on the next page (Common Log Format, Source 2). This is a fragment from the server log files of a so-called “LAMP” server.

- a) Which resource is being requested in line 1?
Does the request succeed? Explain why or why not. (*max 0.5 point*)
- b) Line 2 to 5 suggest there is a problem with this website.
Which one? Motivate your answer. (*max 0.5 point*)
- c) Line 6 and 7 indicate another problem. Describe what is happening on each line. Hint: explain how the “M” and “P” in LAMP relate to the last two lines. (*max 1 point*)

Answer:

Source 2

1	192.16.196.17	--	[08/Apr/2013:14:15:00 +0100]	"GET / HTTP/1.1"	200	2326
2	192.16.196.17	--	[08/Apr/2013:14:15:33 +0100]	"GET /customer_service/ HTTP/1.1"	200	7667
3	192.16.196.17	--	[08/Apr/2013:14:15:38 +0100]	"GET /product_questions/ HTTP/1.1"	200	6324
4	192.16.196.17	--	[08/Apr/2013:14:15:55 +0100]	"GET /customer_service/ HTTP/1.1"	200	7667
5	192.16.196.17	--	[08/Apr/2013:14:15:59 +0100]	"GET /product_questions/ HTTP/1.1"	200	6324
6	192.16.196.17	--	[08/Apr/2013:14:17:59 +0100]	"GET /find_your_id.php?name=John+Doe HTTP/1.1"	200	311
7	192.16.196.17	--	[08/Apr/2013:14:18:39 +0100]	"GET /complaint_form.php?customer_id=0234 HTTP/1.1"	500	32