INF1040 autumn 2006

Compulsory work 1 – Web-pages, text and numbers

Out: Friday 22. September 2006 at 0900 hrs
Due: Friday 6. October 2006 at 1500 hrs

Introduction

The work may be done in small groups (1- 2 students). The answers should be in the form of web-pages in XHTML-strict. The visual appearance of the web-pages should be determined by an external style sheet. You are welcome to use variants of the template on the lecture foils INF1040-CSS-26 to INF1040-CSS-31.

The web-pages should be created by an editor which allows direct editing of the XHTML-text itself. The web-site should be valid under the DTD for XHTML-strict. ***The web-pages should be validated!***

The home page should be named `index.html.` Among other things, this page should tell who has worked out the solutions. This page should also contain the following clause: "I/we have read and understood the rules given in the document "Departemental guidelines for written assignments" on `http://www.ifi.uio.no/studinf/skjemaer/declaration.pdf.` Last, but not least, should the page have links to the solutions to the two assignments below.

You may find a lot of background material on the Internet. You may also use material from the lecture foils and from the course book.Remember to include references whenever you quote or build on material published on the Internet or in print. Good and exact references will be evaluated positively. On the other side: Using other peoples material without giving the source is dishonest and in serious cases a crime, and may in connection with work that contributes to a mark be considered as cheating! See `http://www.ifi.uio.no/studinf/skjemaer/egenerklaring.pdf`

When the job is done, you wrap the files of the web-site into a wrap-file with thje name yourusername_oblig1.zip with the utility program zip, i.e. like this:
zip gerhard_oblig1.zip index.html assignment1.html assignment2.html
Alternatively, you may wrap everything into a tar-file. Then, you send the wrap-file to the group-teacher as attachment to an e-mail. The e-mails should have subject INF1040 – Oblig1. A group of students should send only one file, in that case the username coming fiist in an alphabetical sequence should be used. If the students belong to two different exercise-groups, the file should be sent to the group-teacher who is responsible for the student whose username is first in an alphabetical sequence, whereas the other student sends an e-mail to his group-teacher telling that the answers has been delivered, giving the username of the first student.

Ask the group-teacher if you have any problems!

Assignment 1. XML-files in Unicode

a) Create an XML-document from the following text-paragraph at the bottom of page 26 in Albregtsen og Skagestein: *Digital representasjon*, Unipub forlag 2006:

> Som et eksempel kan vi se på representasjonen i UTF-16 for tegnet ⊕ i det minoiske skriftsystemet Linear B. Unicode-kodepunktet for ⊕ er U+1000F. Vi trekker fra 0x10000 for å få fram kodepunktet relativt til Plan 1: 0x1000F minus 0x10000 gir 0xF = $1111_{2.}$ Vi setter dette bitmønsteret inn som «nyttelast» i «high» og «low surrogate» idet vi begynner bakfra:
>
> High surrogate = **1101 10**00 0000 $0000_{2,}$ hvilket tilsvarer kodepunkt U+D800
> Low surrogate = **1101 11**00 0000 $1111_{2,}$ hvilket tilsvarer kodepunkt U+DC0F

The XML-document should be displayet on the web-page. You may choose element names freely, but they have to give meaning. As character set Unicode in normalform D without precomposed characters should be used. Include in the answer link to the relevant code tables on http://www.unicode.org/

b) Make an estimate of the number of bytes in the XML-document for the Unicode transformations UTF-32, UTF-16 and UTF-8. The estimates should build upon how the characters are represented. (If you want to, you may check that the calculations are approximately correct by storing the XML-documents in the different transformation formats and observe the length, but this should not be the only method!)

In the answer, it may be advantageous to put up a table with a representativ glyph for the special (e.g. the non-ASCII) characters in the paragraph at hand, together with the representations of these characters in the transformations UTF-32, UTF-16 and UTF-8, shown both in binary and in hexadecimal.

Assignment 2. Representation of numbers

In this assignment, you should start with three numbers A, B og C that should be chosen the following way:

- A is a freely chosen integer number between 11 og 44 where the last digit is not greater than 4.
- B is set to –A. Hence, B should be a number between -11 and -44.
- C is set to A+A (or 2*A). Hence, C should be a number between 22 and.

a) Show the representation of the numbers B and C on hexadecimal form in the following formats:
   - Textual form, koded in Unicode UTF-8
   - Binary form (with use of two-complement)
   - Floating point format (IEEE 754)

   In connection with the floating point format, you may use a converter program, for example http://www.h-schmidt.net/FloatApplet/IEEE754.html, but you have to show that the conversion program has given you the correct representations for B and C.

b) Calculate in the binary number system the sum B+C i the three formats, and check that the results always are A. The calculations should be shown on the web-page.
   Hint 1: When you calculate in textual form, you just have to note that B is a negative number and therefore should be subtracted from C.
   Hint 2: If you shift the mantissa of a floating point number on IEEE 754-format one place to the right, the number is halved (remember to include the non-stored bit in the shift-process!) If the value of the exponent is increased by 1, the number is doubled.

Competition!

The best web-sites will receive a prize – a gift certificate for Akademika.

Evaluation criteria:

- The web-site should be usable as teaching material in the subject INF1040
- The technical material presented should be correct
- The web-site should have a good pedagogical structure.

The jury will try to disregard any usage of XHTML-constructions not covered in INF1040, see foil INF1040-XML-25 – this for not giving advantages to competitors who know a lot of XHTML already.

The jury consists of the teachers and the group teachers of INF1040. The decision of the jury is inappealable. The best contribution(s) may be included in the INF1040 teaching material, possibly after some further improvements and tuning.

28.09.2006 GS