XML Technologies: Exercise 5

XSL-FO, XML Parsing

This is a mandatory exercise and the result will be part of your final mark. The solution must be uploaded to OLAT by **April 29th at 15:59**. Late submissions will not be accepted.

Submit the following files in a zipped archive:

- xwing.xsl, xwing.fo, xwing.pdf

- answers.txt

Make sure the archive is named [lastname]_[firstname]_5.zip (for example *mueller_mathias_5.zip*). Some parts of your submission may be automatically evaluated, so make sure to name your files *precisely* as prescribed, otherwise you might not get any points.

1 An XML-Wing Fighter (3 Points)

XSL-FO is a *presentational* markup language. It is written in XML, which has the following advantages:

- XSL-FO documents can easily be generated with XSLT (in fact, that's the most common way to create XSL-FO),

- can easily be manipulated and queried by any XML technology,

- and, most importantly, other markup languages can easily be woven into XSL-FO – in their own namespaces, of course.

In this exercise, we are going to exploit this last fact, namely that XSL-FO easily mixes with other kinds of XML. We will transform content from two different XML documents into XSL-FO and finally render it as PDF.

1.1 Input Documents

The first input documents is `xwing.svg`, which is an SVG file. The purpose of SVG is to provide a means to create graphics that are *scalable*, i.e. that do not have a fixed resolution, very unlike most graphics formats. Therefore, a rendering agent like a web browser can display an SVG graphic in any size.

Check the attached `sample.svg` to get an idea of what a simple SVG file looks like. Pay special notice to attributes such as `fill="#7DD8B5"`, which indicate how elements are coloured.

The second input document `xwing-wiki.xml` is a very simple snippet from Wikipedia.
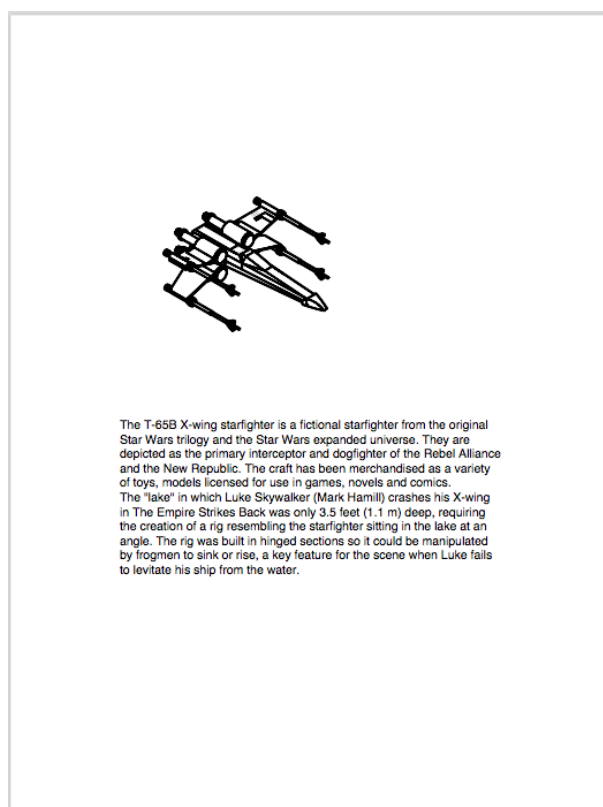
1.2 Transforming Input Documents to XSL-FO and Rendering as PDF

Generally, XSLT is used to generate an FO file, which is then rendered into PDF.

Using the `scaffold.xsl` provided, fill in the blanks to create a FO document, which:

- **First shows the SVG image**

- **Then shows the text for every `<p>`-element in the `xwing-wiki.xml`**

Then render the resulting XSL-FO document as PDF. The resulting PDF should look very similar to the picture below.

Save your XSLT stylesheet as `xwing.xsl`, your XSL-FO file as `xwing.fo` and the rendered PDF as `xwing.pdf`.

2 Parsing XML (2 Points)

XML documents are merely text files stored in the file system. Before an application can access the contents of an XML document, the document must be *parsed*. Roughly speaking, there are two different ways to parse XML – DOM and SAX – and they have fundamental differences.

Answer the following questions:

- **DOM means "document object model". What does that mean? Why is it sometimes called a *tree*?**

- **Which method uses a lot more RAM? Why?**

- **Which of them is an event-driven and streaming method? Why?**

- **Which of them is more suitable for use with XPath? Why? Substantiate your claim with an XPath expression, i.e. find an XPath expression which will not be possible to execute using both approaches.**

Answer in full sentences and save your answers as `answers.txt`.