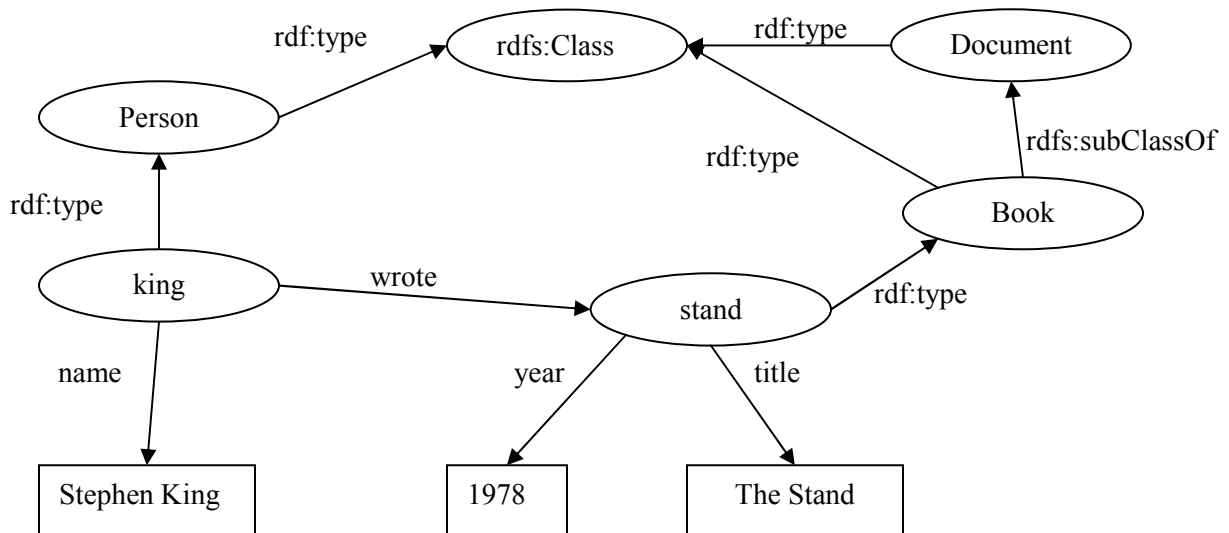


# Homework #1: XML and RDF

The following exercises are due at the beginning of class on Tuesday, September 16. This will count for 10% of your overall grade. Although this homework will be mostly graded for completeness, selected problems will be graded for correctness.

- Determine which of the following XML documents is well-formed. If a document is well-formed, show its corresponding tree model. If it is not well-formed, explain why not. You may assume that all of these documents are intended to be XML 1.0 documents, and thus the declaration is not needed.
  - `<a><b/><a>`
  - `<a><b>foo</b></a><a>bar</a>`
  - `<a><b>foo</b><b>bar</a>`
  - `<a><b><c>foo</b>bar</c></a>`
  - `<a/><b>foo</b><b>bar<b>`
  - `<a><b><c>foo</c>baz<c>bar</c></b></a>`
  - `<a x="1" X="2"><b>foo</b></a>`
  - `<a x="1"><b x="2">foo</b></a>`
  - `<a x="1" y="2" x="3"><b>foo</b></a>`
  - `<a><b x="3"><c y="1">foo</c><c>bar</c></b></a>`
- Translate the following RDF Graph into the RDF XML syntax. Be careful to give the syntax that exactly translates into the provided graph. In particular, do not include syntax for any implied or inferred triples. Assume that rectangular nodes represent untyped literals. For nodes and arcs labeled with QNames, assume the standard prefixes apply. For other names, assume they are all local to the document you are writing.



3. Consider the following RDF document using the XML syntax. Draw the equivalent graph. For convenience, you may use QNames for your node and edge labels.

```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:u="http://www.example.org/uni#"
  xml:base="http://www.example.org/uni">
<rdfs:Class rdf:ID="Person" />
<rdfs:Class rdf:ID="Student" />
  <rdfs:subClassOf="#Person" />
</rdfs:Class>
<rdfs:Class rdf:ID="Professor">
  <rdfs:subClassOf="#Person" />
</rdfs:Class>
<rdfs:Class rdf:ID="Course" />
<rdf:Property rdf:ID="advises">
  <rdfs:domain rdf:resource="#Professor" />
  <rdfs:range rdf:resource="#Student" />
  <rdfs:subPropertyOf="#knows">
</rdf:Property>
<rdf:Property rdf:ID="takes">
  <rdfs:domain rdf:resource="#Student" />
  <rdfs:range rdf:resource="#Course" />
</rdf:Property>
<rdf:Property rdf:ID="teaches">
  <rdfs:domain rdf:resource="#Professor" />
  <rdfs:range rdf:resource="#Course" />
</rdf:Property>
<rdf:Property rdf:ID="knows" />
<u:Professor rdf:ID="alan">
  <u:teaches rdf:resource="#cs100" />
  <u:advises rdf:resource="#rob" />
  <u:advises rdf:resource="#sarah" />
</u:Professor>
<u:Student rdf:ID="rob">
  <u:takes rdf:resource="#cs100" />
  <u:takes rdf:resource="#cs200" />
</u:Student>
</rdf:RDF>
```

4. Using the RDFS entailment rules `rdfs2`, `rdfs3`, `rdfs5`, `rdfs7`, `rdfs9`, and `rdfs11` (see the RDF Semantics recommendation (<http://www.w3.org/TR/rdf-mt/>), Section 7), determine what triples can be inferred from the example in Problem #3. You may ignore the other entailment rules because they do not add anything particularly interesting. You may give your answer either via extending the graph (but use a different line color or style to indicate new triples) or by simply listing the triples.
5. According to the RDFS Schema recommendation, a property can have multiple `rdfs:range` statements. When this occurs the range of the property is treated as the intersection of the individual ranges. What advantage does this have over using the union of the individual ranges to determine the actual range? Hint: Consider the distributed design of RDF and the inferences that are sanctioned by `rdfs:range`.