


<oo> → <dh> Digital humanities

Maintained by: David J. Birnbaum (djbitt@gmail.com) 

Last modified: 2013-03-04T17:46:12+0000

Schematron assignment #1

In a three-way election for Best Stooge Ever, each candidate (Curly, Larry, Moe) wins between 0% and 100% of the votes. Assume that all votes are cast for one of the three candidates (no abstentions, write-ins, invalid ballots, etc.), which means that when you add the percentages for the three candidates, the result must be exactly 100%. Assume also that we're recording percentage of the vote, not raw votes, and that the percentages are all integer values. (In Real Life we'd probably record the raw count and *calculate* the percentages, but in real life we wouldn't be voting for Best Stooge Ever in the first place!) Here's a Relax NG schema for the results of the election:

```
start = results
results = element results { stooge+ }
stooge = element stooge { name, xsd:int }
name = attribute name { "Curly" | "Larry" | "Moe" }
```

Here's a sample XML document that is valid against the preceding schema:

```
<results>
  <stooge name="Curly">50</stooge>
  <stooge name="Larry">35</stooge>
  <stooge name="Moe">15</stooge>
</results>
```

We could have written a better Relax NG schema, but we didn't, and although our sloppy schema works with the results above, it also allows erroneous results like the following:

```
<results>
  <stooge name="Curly">55</stooge>
  <stooge name="Larry">38</stooge>
  <stooge name="Moe">11</stooge>
</results>
```

The problem here is that the three "percentage" values total 104%, and no matter how good our coding, it is not possible to prevent this type of error by using Relax NG alone. Your assignment is to write a Schematron schema that verifies that the three percentages always total exactly 100%. Test your results by creating the Relax NG schema, your Schematron schema, and a sample XML document that you can validate against both schemas in <oXygen/>. Enter correct and incorrect values and verify that the Schematron schema is working correctly. For homework, upload only your Schematron schema.

You can stop here and consider the assignment complete, but for more Schematron practice, you're welcome to add additional rules to check for additional types of error. The following types of errors could have been controlled by writing a better Relax NG schema, but for the purpose of learning Schematron, let's do it in Schematron:

1. There should be exactly three votes, with exactly one for each Stooge. No duplicate Stooges and no missing Stooges.
2. Each individual Stooge's vote should range from 0 to 100. No negative integers and no integers greater than 100. (The Relax NG schema is ensuring that all values are integers, so you don't have to worry about that.)