A semantic markup language for curating physics content on the web

Ian LAWRENCE,

Institute of Physics, 80 Portland Place, London W1B 1NT, UK

Expressing physics on a web page with the fruitful precision required in order to best support learning tends to be a demanding technical task for authors and content curators. That is as as a result of the need to adhere to various conventions, such as the spacing between numbers and units, and ensure that unhelpful distributions of words do not detract from meaning, such as splitting equations over lines, or placing number and unit on different lines, whatever the display device. In short, there is a lot to get right. Current standards in markup languages such as html do not lend much support to avoiding of these kinds of mis-steps. Some current techniques suggest a mixture of technologies, such as embedding LATeX via MathJax to represent equations, but this is inelegant, introduces dissipative friction into the authoring and editing systems, and results in extra difficulties in separating content from display technologies whilst ensuring representational consistency across the different technologies deployed.

The poster will present a designed and implemented lightweight semantic markup language, optimised for publishing and curating physics content. It has roots in both significant use of ConTeXt, a TEX macro language, and MarkDown. A system using this physics markup language can be implemented using standard web technologies (generating clean, standards-compliant and mobile-friendly html, with hooks for CSS, and relying only on JavaScript).

Authoring can be in any text editor, with all the power that this implies, but can also be in a dual-pane, browser-based editor. In either case significant authorial and editorial support can be provided by use of custom auto-completion and code colouring. Authoring or editing can however, take place in any text editor, however simple, lowering the technical barriers for potential contributors.

A version of this system, using a python script to generate the html, has been used to curate and publish the 400,000 plus words of the Supporting Physics Teaching Project(http://supportingphysicsteaching.net/) for the last five years. An evolved variant is now the editing core of the upcoming online teacher centre from the Institute of Physics. So the system is well tested for representing the non-diagrammatic elements physics for 5-19 year old children.