

Grammar for Mark's File Format

COSC-575

January 30, 2012

Grammar

$\langle dataset \rangle \rightarrow \langle header \rangle \langle attributes \rangle \langle examples \rangle$
 $\langle header \rangle \rightarrow '@dataset' \langle identifier \rangle$
 $\langle attributes \rangle \rightarrow \langle attribute \rangle \langle attribute \rangle \langle attributes \rangle \mid \epsilon$
 $\langle attribute \rangle \rightarrow '@attribute' \langle declaration \rangle$
 $\langle declaration \rangle \rightarrow 'numeric' \langle identifier \rangle \mid$
 $\quad 'discrete' \langle identifier \rangle \langle levels \rangle \mid$
 $\quad 'nominal' \langle identifier \rangle \langle nominal-values \rangle$
 $\langle examples \rangle \rightarrow '@examples' \langle example-list \rangle$
 $\langle example-list \rangle \rightarrow \langle example \rangle \langle examples-list \rangle \mid \epsilon$
 $\langle examples \rangle \rightarrow \langle attribute-value \rangle \langle example \rangle \mid \epsilon$
 $\langle attribute-value \rangle \rightarrow \langle identifier \rangle \mid \langle number \rangle$
 $\langle nominal-values \rangle \rightarrow \langle identifier \rangle \langle identifier \rangle \langle identifier-list \rangle$
 $\langle identifier-list \rangle \rightarrow \langle identifier \rangle \langle identifier-list \rangle \mid \epsilon$
 $\langle identifier \rangle \rightarrow \langle non-whitespace-character \rangle \langle identifier \rangle \mid \epsilon$
 $\langle levels \rangle \rightarrow '1', '2', '3', \dots$
 $\langle number \rangle \rightarrow -\infty, \dots, +\infty$
 $\langle non-whitespace-character \rangle \rightarrow 'a', \dots, 'z', 'A', \dots, 'Z', '0', \dots, '9', '-', \dots, '+'$