The Welkin Standard

Syntax

Terminals

- Logic
- Undefined notions:
  - Symbols: $0, 1$
  - Successor S
  - Implication $\Rightarrow$
- Table of US-ASCII:

Symbol	Encoding
$\{$	173
$\}$	175

- Empty set: ε satisfies $\neg S(\varepsilon)$
- A **word** is recursively defined.
  - Base case: the empty set ε is a word.
  - Recursion: let w be a word.
    - $w.0$ is a word.
    - $w.1$ is a word.
- Concatenation
  - Base case: $w.\varepsilon = w$.
  - Recursion:
    - $w.(u.0) = (w.u).0$
    - $w.(u.1) = (w.u).1$

Atoms

- Strings are words with delimiters: $d_1.\text{w}.d_2$, where $d_1 \not\subset w$ and $d_2 \not\subset w$.
- Identifiers are strings without white space.
- Numbers are a subset of strings with an injective function $q : \text{NUMBER} \to Q$.
  - Q is set of strings formed by scientific notation.

Grammar

- LALR
  - Not ambiguous
- Welkin Grammar:

Semantics

Equality on Terms

- Two strings are equal if they contain the same strings, in order.
- Two numbers are equal if $q(a) = q(b)$.

Valid Strings

- No relative members at toplevel (with length 2).

- No duplicate members, graphs, or connections.

Welkin Information Graphs

A **Welkin Information Graph (WIG)** is a structure $G = (T, H, L)$ with:

- A tree T,
- A hypergraph H,
- A tree L isomorphic to T.

AST (Recursive)

- Units:
- Members are words of units
- Connections are WIGs with
- Graphs are WIGs with
  - Derived terms as children
  - Ordered triples are arcs.

Encoding

The **encoding** $E(G)$ of the WIG G is the unique string where

- All nodes are listed in breadth-first order
- Leaves are terms ending with "#"
- Edges are enumerated, starting from 0. They are included in nodes:
  - s means source,
  - c means connector,
  - t means target.