unique readability of well-formed formulas

Canonical name	UniqueReadabilityOfWellformedFormulas
Date of creation	2013-03-22 18:52:32
Last modified on	2013-03-22 18:52:32
Owner	CWoo (3771)
Last modified by	CWoo (3771)
Numerical id	9
Author	CWoo (3771)
Entry type	Theorem
Classification	msc 03B05
Defines	unique readability

Suppose V is an alphabet. Two words w_1 and w_2 over V are the same if they have the same length and same symbol for every position of the word. In other words, if $w_1 = a_1 \cdots a_n$ and $w_2 = b_1 \cdots b_m$, where each a_i and b_i are symbols in V, then $n = m$ and $a_i = b_i$. In other words, every word over V has a unique representation as product (concatenation) of symbols in V. This is called the *unique readability* of words over an alphabet.

Unique readability remains true if V is infinite. Now, suppose V is a set of propositional variables, and suppose we have two well-formed formulas (wffs) $p := \alpha p_1 \cdots p_n$ and $q := \beta q_1 \cdots q_m$ over V, where α, β are logical connectives from a fixed set F of connectives (for example, F could be the set $\{\neg, \vee, \wedge, \to, \leftrightarrow\}$), and p_i, q_j are wffs. Does $p = q$ mean $\alpha = \beta$, $m = n$, and $p_i = q_i$? This is the notion of *unique readability* of well-formed formulas. It is slightly different from the unique readability of words over an alphabet, for p_i and q_j are no longer symbols in an alphabet, but words themselves.

Theorem 1. *Given any countable set V of propositional variables and a set F of logical connectives, well-formed formulas over V constructed via F are uniquely readable*

Proof. Every $p \in \overline{V}$ has a representation $\alpha p_1 \cdots p_n$ for some n-ary connective α and wffs p_i. The rest of the proof we show that this representation is unique, establishing unique readability.

Define a function $\phi : V \cup F \to \mathbb{Z}$ such that $\phi(v) = 1$ for any $v \in V$, and $\phi(f) = 1 - n$ for any n-ary connective $f \in F$. Defined inductively on the length of words over $V \cup F$, ϕ can be extended to an integer-valued function ϕ^* on the set of all words on $V \cup F$ so that $\phi^*(w_1 w_2) = \phi^*(w_1) + \phi^*(w_2)$, for any words w_1, w_2 over $V \cup F$. We have

1. ϕ^* is constant (whose value is 1) when restricted to $\overline{V}$.

   This can be proved by induction on V_i (for the definition of V_i, see the parent entry). By definition, this is true for any atoms. Assume this is true for V_i. Pick any $p \in V_{i+1}$. Then $p = \alpha p_1 \cdots p_n$ for some n-ary $\alpha \in F$, and $p_j \in V_i$ for all $j = 1, \ldots, n$ Then $\phi^*(p) = \phi^*(\alpha) + \phi^*(p_1) + \cdots + \phi^*(p_n) = 1 - n + n \times 1 = 1$. Therefore, $\phi^*(p) = 1$ for all $p \in \overline{V}$.

2. for any non-trivial suffix s of a wff p, $\phi^*(s) > 0$. (A *suffix* of a word w is a word s such that $w = ts$ for some word t; s is non-trivial if s is not the empty word)

This is also proved by induction. If $p \in V_0$, then p itself is its only non-trivial final segment, so the assertion is true. Suppose now this is true for any proposition in V_i. If $p \in V_{i+1}$, then $p = \alpha p_1 \cdots p_n$, where each $p_k \in V_i$. A non-trivial final segment s of p is either p, a final segment of p_n, or has the form $tp_j \cdots p_n$, where t is a non-trivial final segment of p_{j-1}. In the first case, $\phi^*(s) = \phi^*(p) = 1$. In the second case, $\phi^*(s) > 0$ from assumption. In the last case, $\phi^*(s) = \phi^*(t) + (n - j + 1) > 0$.

Now, back to the main proof. Suppose $p = q$. If p is an atom, so must q, and we are done. Otherwise, assume $p = \alpha p_1 \cdots p_m = \beta q_1 \cdots q_n = q$. Then $\alpha = \beta$ since the expressions are words over $V \cup F$, and $\alpha, \beta \in F$. Since the two connectives are the same, they have the same arity: $m = n$, and we have $\alpha p_1 \cdots p_n = \alpha q_1 \cdots q_n$. If $n = 0$, then we are done. So assume $n > 0$. Then $p_1 \cdots p_n = q_1 \cdots q_n$. We want to show that $p_n = q_n$, and therefore $p_1 \cdots p_{n-1} = q_1 \cdots q_{n-1}$, and by induction $p_i = q_i$ for all $i < n$ as well, proving the theorem.

First, notice that p_n is a non-trivial suffix of $q_1 \cdots q_n$. So p_n is either a suffix of q_n or has the form $tq_j \cdots q_n$, where $j \leq n$, and t is a non-trivial suffix of q_{j-1}. In the latter case, $1 = \phi^*(p_n) = \phi^*(tq_j \cdots q_n) = \phi^*(t) + n - j + 1$. Then $\phi^*(t) = j - n \leq 0$, contradicting 2 above. Therefore p_n is a suffix of q_n. By symmetry, q_n is also a suffix of p_n, hence $p_n = q_n$. $\square$

As a corollary, we see that the well-formed formulas of the classical propositional logic, written in Polish notation, are uniquely readable. The unique readability of wffs using parentheses and infix notation requires a different proof.

Remark. Unique readability will fail if the p_i and q_j above are not wffs, even if V is finite. For example, suppose $v \in V$ and $\#$ is binary, then $\#v\#vv$ can be read in three non-trivial ways: combining v and $\#vv$, combining $v\#$ and vv, or combining $v\#v$ and v. Notice that only the first combination do we get wffs.