Hall’s Theorem

The set of propositional connectives is $\mathcal{C} = \{\neg, \lor, \land, \rightarrow, \leftrightarrow\}$. The arities of these connectives are 1 for negation $\neg$, and 2 for all the others. The set of propositional variables is $\mathcal{P} = \{p, q, r, \ldots\}$. A word is a non-empty string of symbols taken from $\mathcal{C} \cup \mathcal{P}$. Certain words are formulas, $\lor pq$, others are not, e.g., $p \lor q$. Note, that we are going to use Polish notation which is due to J. Łukasiewicz.

The following is an algorithm due to P. Hall. First, we define the length $l(w)$ of a word $w$ as the number of characters in it. E.g., if $w = x_1x_2 \ldots x_n$ then $l(w) = n$. Thus, $l(ww') = l(w) + l(w')$.

Next, we define the valency $v(w)$ of a word $w$. There are two types of words of length one, connectives and variables. If $c$ is a connective then we define $v(c) = (\neg n+1)$ where $n$ is the arity of $c$. Thus, $v(\neg) = 0$ and $v(c) = -1$ for all the other connectives. If $p$ is a variable then we define $v(p) = 1$. Then, if $w = x_1 \ldots x_n$, we define $v(w) = v(x_1) + \ldots + v(x_n)$. Again, $v(ww') = v(w) + v(w')$.

For example, $v(\lor pq) = v(\lor) + v(p) + v(q) = (-2 + 1) + 1 + 1 = 1$.

We may think that we input from right to left the variables $q$ and $p$ which yields valency 2 and then the $\lor$ absorbs the two variables yielding a proper formula of valency 1.

**Theorem.** The word $w$ is a product of $k$ formulas, i.e.,

$$w = w_1 \ldots w_k$$

if and only if the Hall condition holds:

$$v(w) = k \text{ and } v(u) > 0 \text{ for every right segment } u \text{ of } w.$$

**Corollary.** A word $w$ is a formula if and only if $v(w) = 1$ and $v(u) > 0$ for every right segment $u$ of $w$.

**Proof.** We are proceeding by induction on the length of $w$. Assume that $w$ is a word of length $n$ and that for all words of length $< n$ the Hall condition is equivalent to that $w$ is a product (i.e., concatenation) of propositional formulas.

Assume first that the Hall condition holds for $w$. We have

$$w = x_1 \ldots x_n = x_1w'$$

where $x_1$ is either a connective $c$ or a propositional variable $p$. At any rate, $w'$ satisfies the Hall condition: Every right segment $u$ of $w'$ is a right segment of $w$ and therefore $v(u) > 0$. Thus, $w'$ is a product of words:

$$w' = w_1 \ldots w_m$$

where $v(w') = m$. Now, if $x_1$ is a propositional variable $p$, then

$$w = pw' = pw_1 \ldots w_m$$

shows that $w$ is a product of $m+1$ formulas where $v(w) = v(p) + v(w') = 1 + m$. The other possibility is $x_1 = c$ for one of the connectives $c$ in $\mathcal{C}$. Assume first that $c = \neg$. Then $v(w) = v(c) + v(w') = 0 + m = m$ and

$$w = (\neg w_1) \ldots w_m$$

is a product of $m$—many formulas. If $c$ is one of the binary connectives then $v(w) = -1 + m > 0$ shows that $m \geq 2$ and

$$w = (cw_1w_2) \ldots w_m$$

is a product of $v(w) = m - 1$ formulas. We have shown the more interesting part of the Theorem: Hall’s condition is a test for that a string of characters is well-formed, i.e., a concatenation of formulas. Now assume that we are given a product $w = w_1 \ldots w_m$ of formulas. We wish to show that Hall’s condition holds. Again we are proceeding by induction on the length of $w$. If the first character of $w$ is a propositional variable $p$ then it must be the case that $w_1 = p$ and we let $w' = w_2 \ldots w_m$. By induction hypothesis we have that for all right segments $u$ of $w'$ that $v(u) > 0$ and $v(w') = m - 1$. The right segments of $w$ are those of $w'$ and $w$ itself. But $v(w) = m - 1 + 1 = m$ and we are done in this case.
Now assume that $w_1$ is a complex formula, say $w_1 = \land \alpha \beta$. Then
\[ w'' = \alpha \beta w_2 \ldots w_m \]
is a product of $m + 1$ formulas but shorter than $w$. By induction, for all of its right segments $u$ we have that $v(u) > 0$ and $v(w'') = m + 1$. We have that these are the right segments of $w$ plus $w$ itself. But $v(w) = (-1) + v(w'') = m$ and we are done. \qed

We also need the

**Formation Lemma.** Assume that $w$ is a product of formulas $w_i$, $w'_j$:
\[ w = w_1 \ldots w_n = w'_1 \ldots w'_m \]
then $m = n$ and $w_i = w'_i$.

**Proof.** We proceed again by induction on the length of $w$. The case $l(w) = 1$ is, of course, trivial: $w = p$. Assume for all words of length $< n$ the following: If $w$ is a product of formulas,
\[ w = w_1 \ldots w_k = w'_1 \ldots w'_l \]
then $k = l$ and $w_i = w'_i$.
If the first character of $w$ is a propositional variable $p$ then $p = w_1$. There is no formula of length $> 1$ that starts with a $p$. So $p = w_1 = w'_1$ and we can cancel $p$ on both sides and apply induction.
If $w_1$ is, for example, a conjunction, say $w_1 = \land \alpha \beta$ then $w = (\land \alpha \beta) w_2 \ldots w_k$ then also $w'_1$ must be a conjunction, thus
\[ w = \land \alpha \beta w_2 \ldots w_k = \land \alpha' \beta' w'_2 \ldots w'_l \]
Thus
\[ \alpha \beta w_2 \ldots w_k = \alpha' \beta' w'_2 \ldots w'_l \]
and we conclude, by induction, that $\alpha = \alpha'$, $\beta = \beta'$, $k = l$, $w_i = w'_i$. But then also $w_1 = w'_1$. \qed