A PROOF OF ISBELL’S ZIGZAG THEOREM

PIOTR HOFFMAN

(Received 25 January 2006; accepted 1 August 2006)

Communicated by D. Easdown

Abstract

We provide a short, intuitive proof of Isbell’s zigzag theorem.


Keywords and phrases: semigroups, zigzags.

The elegant, elementary proof of the zigzag theorem published by Higgins [1, 2] is unfortunately incorrect (the case overlooked in the proof is when the transition \( \alpha \) is \( puq \to ptutq \) and \( \beta \) is \( ptutq \to puqstq \), where \( u = u's \) in \( S \)). The proofs of Howie [3], Storrer [4] and Renshaw [5] involve nontrivial algebra. The proof of Philip [6], which completes Isbell’s original proof [7], is topological in nature. We provide a short, elementary and intuitive proof of the monoid version of the zigzag theorem. Our proof is heavily based on the proposed proof of Higgins.

Recall [8] that if \( U \) is a submonoid of \( S \), then the dominion of \( U \) in \( S \) is the set of all elements \( s \in S \) such that, for all homomorphisms \( h_1, h_2 : S \to T \) coinciding on \( U \), \( h_1(s) = h_2(s) \). Note that a homomorphism \( h : U \to S \) is epi if and only if the dominion of \( h(U) \) in \( S \) is \( S \).

**Isbell’s Zigzag Theorem.** Let \( U \) be a submonoid of a monoid \( S \) and \( s_* \) an element of \( S \). Then \( s_* \) is in the dominion of \( U \) in \( S \) if and only if there exists a \( U \)-zigzag in \( S \) with value \( s_* \), that is, a diagram of the form

\[
\begin{array}{cccccccc}
& & & & & & & \\
& u_0 & & u_1 & & u_2 & & \ldots & u_n & u_{n+1} \in U \\
& s & & s_1 & & s_2 & & \ldots & s & s_* \\
\end{array}
\]

where \( s_1, \ldots, s_n, s_1, \ldots, s \in S, u_0, u_1, u_2, \ldots, u_n, u_{n+1} \in U \), and all cells in the diagram commute, that is, \( s_* = s_1u_1^0, u_1^0 = u_1^1s_1, s_2u_2^1 = u_2^2s_2^1, \ldots, s_nu_n^n = u_{n+1}^n, u_{n+1}^nS = s_* \).

(Received 25 January 2006; accepted 1 August 2006)
A simple verification shows that indeed, if $h_1, h_2 : S \to T$ are homomorphisms coinciding on $U$ and there is a $U$-zigzag with value $s_*$, then $h_1(s_*) = h_2(s_*)$, which proves that $s_*$ is in the dominion.

So, assume now that $s_*$ is in the dominion of $U$ in $S$.

Let $A$ consist of all elements $s$ of the monoid $S$ and of a new element $\varepsilon$. Consider the set $A^*$ of all finite words over the alphabet $A$, with $\varepsilon$ denoting the empty word. On $A^*$, define three types of reductions:

- **refactorization** $s_1 \ldots s_n \leftrightarrow s'_1 \ldots s'_k$ if $s_1 \cdot \ldots \cdot s_n = s'_1 \cdot \ldots \cdot s'_k$ holds in $S$ ($n, k \geq 0$ and $s_1, \ldots, s_n, s'_1, \ldots, s'_k \in S$);

- **right/left shift** $|u \leftrightarrow u|$ for $u \in U$;

- **creation/deletion** $\varepsilon \leftrightarrow ||$.

Let $\leftrightarrow$ be the reduction relation defined by the above reductions, and let $\leftrightarrow^+$ be its transitive closure; then $\leftrightarrow^+$ is a congruence on $A^*$, giving rise to a quotient monoid $T$.

Consider maps $\mu, \nu : S \to T$ given by $\mu(s) = s$ and $\nu(s) = |s|$ for all $s \in S$. Both $\mu$ and $\nu$ are monoid homomorphisms, and they coincide on $U$. Therefore we have $\mu(s_*) = \nu(s_*)$. In other words, $s_* \leftrightarrow^+ |s_*|$, which is equivalent to $s_* \leftrightarrow^+ |s_*|$. We will show that if a sequence $q$ of reductions

$$s_* = w_1 \leftrightarrow w_2 \leftrightarrow \ldots \leftrightarrow w_{n-1} \leftrightarrow w_n = |s_*|$$

exists ($n \geq 2$), then there is a $U$-zigzag in $S$ with value $s_*$.

The proof rests on two observations.

The first observation is that one may track particular occurrences of the symbol $|$. That is, any occurrence of $|$, if it is not the occurrence appearing in $w_1$, is first created by some reduction $w_{i-1} \leftrightarrow w_i$; then $i$ is called its birth and the occurrence born with it is called its birth pair. The identity of an occurrence of $|$ may be tracked along any shifts it performs, as well as any reductions in which it does not take an active role. If the tracked occurrence is not the one appearing in $w_n$, then it must be deleted by some reduction $w_j \leftrightarrow w_{j+1}$; $j$ is then called its death and the occurrence that died with it is called its death pair. Potential problems that could arise in reductions, such as $| \leftrightarrow ||$, where it is not clear whether the newly-born pair of occurrences appears to the right or
to the left of the original occurrence, can be resolved by, for instance, always choosing births to happen at the left end of a word; the same remark applies to deaths.

The second observation is that if \( w_i \leftrightarrow w_{i+1} \), and if \( w_i \) contains an occurrence of \( \mid \) that does not die at \( i \), then also \( v_i \leftrightarrow v_{i+1} \), where \( v_i \) and \( v_{i+1} \) arise from \( w_i \) and \( w_{i+1} \) by removing all occurrences of \( \mid \) other than the one under consideration. The reason is that a refactorization or a shift of the considered occurrence of \( \mid \) remain uninfluenced, while creations, deletions or shifts involving occurrences other than the considered one are replaced by trivial refactorizations.

Consider the following sequence \( r \) of reductions: start with \( w_1 \), and let \( o_1 \) be the original occurrence of \( \mid \) in \( w_1 \). Proceed as in \( q \) up to the point \( i_1 \) of \( o_1 \)'s death, removing all occurrences of \( \mid \) other than \( o_1 \); by the second observation, this leads to a correct reduction sequence. Let \( o_2 \) be the death pair of \( o_1 \). Observe that \( w_{i_1} \) with occurrences other than \( o_1 \) removed is the same as \( w_{i_1} \) with occurrences other than \( o_2 \) removed, since \( o_1 \) and \( o_2 \) are adjacent in \( w_{i_1} \). So we may now proceed as in \( q \), but in reverse, from \( i_1 \) to the point \( i_2 \) of \( o_2 \)'s birth, removing all occurrences of \( \mid \) other than \( o_2 \). Let \( o_3 \) be the birth pair of \( o_2 \). The above procedure may be repeated until an occurrence \( o_m \) appears that does not die at all; then \( o_m \) must coincide with the sole occurrence of \( \mid \) in \( w_n \). By this procedure, depicted in Figure 1, we have built a sequence of reductions \( r \) starting with \( s_\ast \mid = w_1 \), ending with \( w_n = \mid s_\ast \), and containing only words in which \( \mid \) appears exactly once; thus, \( r \) consists of refactorizations and shifts only.

For any word \( w = s_1 \ldots s_n |s'_1 \ldots s'_k \) in the sequence \( r \), consider the three-letter word \( w' = (s_1 \ldots s_n)(s'_1 \ldots s'_k) \). This gives rise to a sequence \( r' \), which starts with \( s_\ast |1 \) and ends with \( 1|s_\ast \). If \( w \leftrightarrow v \) was a refactorization in \( r \), then \( w' = v' \). If \( w \leftrightarrow v \) was a right shift in \( r \) involving \( u \in U \), then \( w' = s|(u \cdot s') \) and \( v' = (s \cdot u)|s' \) for some \( s, s' \in \mathcal{S} \); similarly for left shifts. Thus, \( r' \) precisely corresponds to a \( U \)-zigzag in \( \mathcal{S} \) with value \( s_\ast \). This completes the proof.

References


PIOTR HOFFMAN, Warsaw University, Institute of Informatics, ul. Banacha 2, 02-097 Warszawa, Poland

e-mail: piotrek@mimuw.edu.pl