

For any presentation ϕ and time t , define k -SPIA (Strictly k -Piecewise Inference Algorithm) as follows

$$k\text{-SPIA}(\phi\langle t \rangle) = \begin{cases} \emptyset & \text{if } t = 0 \\ k\text{-SPIA}(\phi\langle t-1 \rangle) \cup \text{subseq}_k(\phi(t)) & \text{otherwise} \end{cases}$$

Note that we are being a little sloppy here. Technically, the output of k -SPIA given some input sequence is a set of subsequences G , not a program. What we really mean with the above is that k -SPIA outputs a program which uses G to solve the membership problem for $L(G) = \{w \mid \text{subseq}_k(w) \subseteq G\}$. This program looks something like this.

1. Input: any word w .
2. Check whether $\text{subseq}_k(w) \subseteq G$.
3. If so, OUTPUT Yes, otherwise OUTPUT No.

All k -SPIA does is update this program simply by updating the contents of G .

Theorem 1. *For each k , k -SPIA identifies in the limit from positive data the collection of stringsets SP_k .*

Proof Consider any $k \in \mathbb{N}$. Consider any $S \in SP_k$. Consider any positive presentation ϕ for S . It is sufficient to show there exists a point in time t_ℓ such that for all $m \geq t_\ell$ the following holds:

1. $k\text{-SPIA}(\langle m \rangle) = k\text{-SPIA}(\langle t_\ell \rangle)$ (convergence), and
2. $k\text{-SPIA}(\langle m \rangle)$ is a program that solves the membership problem for S .

Since $S \in SP_k$, there is a finite set $G \subseteq \Sigma^{\leq k}$ such that $S = L(G)$.

Consider any subsequence $g \in G$. Since $g \in G$ there is some word $w \in S$ which contains g as a k -subsequence. Since G is finite, there are finitely many such w , one for each g in G . Because ϕ is a positive presentation for S , there is a time t where each of these w occurs. For each w let t be the first occurrence of w in ϕ . Let t_ℓ denote the latest time point of all of these time points t . Next we argue that for all time points m larger than this t_ℓ , the output of k -SPIA correctly solves the membership problem for S and does not change.

Consider any $m \geq t_\ell$. The claim is that $k\text{-SPIA}(\langle m \rangle) = k\text{-SPIA}(\langle t_\ell \rangle) = G$. For each g in G , a word containing g as a subsequence occurs at or earlier than t_ℓ and so $g \in k\text{-SPIA}(\langle m \rangle)$. Since g was arbitrary in G , $G \subseteq k\text{-SPIA}(\langle m \rangle)$.

Similarly, for each $g \in k\text{-SPIA}(\langle m \rangle)$, there was some word w in ϕ such that w contains g as a subsequence. Since ϕ is a positive presentation for S , w is in S . Since w belongs to S , $\text{subseq}_k(w) \subseteq G$ and so g belongs to G . Since g was arbitrary in $k\text{-SPIA}(\langle m \rangle)$ it follows that $k\text{-SPIA}(\langle m \rangle) \subseteq G$.

It follows $k\text{-SPIA}(\langle m \rangle) = G$.

Since m was arbitrarily larger than t_ℓ we have both convergence and correctness.

Since ϕ was arbitrary for S , S arbitrary in SP_k and k arbitrary, the proof is concluded. \square