Word Equations: Sheet 5

Task 1 Improve Rytter's algorithm and/or its analysis so that the approximation ratio is $\mathcal{O}(\log(n/g))$, where g is the size of the smallest grammar for the string. It is also enough to show that the size of the produced grammar is of size $\mathcal{O}(\ell \log(n/\ell))$, where ℓ is the size of the LZ77 factorisation.
Hint: You can either subdivide the initial text into smaller fragments and at the end concatenate them or make the analysis tighter, by showing that f_i can be concatenated in $\mathcal{O}(\log |f_i|)$, and estimating the sum appropriately.

Task 2 Improve Rytter's algorithm and/or its analysis so that it can handle self-referencing LZ77. What is the obtained approximation ratio?

Task 3 Using Rytter's approach give an algorithm for transforming the composition system into an SLP. What is the size-increase?

Task 4 Addition chains An *addition chain* for numbers $T = \{n_1, n_2, n_m\}$ is a sequence of numbers $n'_1, n'_2, \ldots, n'_{m'}$ such that each of them is either 1 or a sum of two previous numbers; furthermore, each number n_i is an element of $n'_1, n'_2, \ldots, n'_{m'}$. The size of the addition chain is m'.

Define the string $w = a^{n_1}\$a^{n_2}\$\cdots\$a^{n_m}$. Let m' be the length of the shortest addition chain for T and let g the size of the smallest grammar for w. Show that $m' = \Theta(g)$.

Task 5 Based on a paper, 2 points Show that for $T = \{n_1, n_2, \ldots, n_m\}$ one can construct an addition chain of size $\lg(\sum_i n_i) + \mathcal{O}\left(\sum_i \frac{\lg n_i}{\lg \lg n_i}\right)$

Task 6 Show the NP-hardness of the decision variant of the for smallest grammar problem, i.e. given a string w and number m decide, whether there is an SLP for w of size at most m. Show the inapproximability result: there is a constant α such that existence of an α-approximation algorithm implies that P=NP.
Hint: You can alter the definition of the SLP, so that the size is calculated somehow different or that rules are of slightly more general form.