Alice in Wonderland—The tale of the long tail in globally optimized page breaking (part 2)

Frank Mittelbach
\LaTeX3 Project, Mainz, Germany

July 15, 2016

The story of Alice's Adventures In Wonderland by Lewis Carroll contains a long tale about the tail of the mouse which starts like

Fair said to a mouse, That he met in the house,
"Let us both go to law: I will prosecute YOU.---Come, I'll take no denial; We must have a trial: For really this morning I've nothing to do."

Said the mouse to the cur, "Such a trial, dear Sir, ...

... many more lines ...

Obviously a tail like this should be typeset in full beauty and should not be laid out in knots or cut between pages. Unfortunately, that is more easily said
than done, given that all typesetting systems use a greedy page algorithm that
cuts page by page. Thus chances are high that disaster strikes and we have to
manually adjust earlier page breaks to prevent it.

Using global optimization in pagination has been envisioned already more
than 30 years ago by Michael Plass in his PhD thesis and throughout the years
other people worked on specific aspects of global optimized pagination, but
until today all typesetting engines have taken the “easy” way out and leave the
problem essentially to the user.

This is in fact not that surprising, as the problem can get easily out of hand:
a naive approach will immediately result in exponential complexity (without
introducing float objects into the mix which makes it worse). But even with
careful restrictions and specialized algorithms we will soon be back at testing
the computer limits with any real live document.

However, computers are getting faster and thus get us closer to make globally
optimized pagination a reality. So this year I started to implement a framework
that assists in this task, in the hope that for my next book I do not have to
hand-adjust half of the page breaks manually—first results are promising!

At Bachotek 2016 I gave some theoretical background to the problem and
discussed the basic approach taken by the framework, including two already
implemented optimization strategies: The automatic change of page length on
double spreads to add flexibility and the use of automatic variation in paragraph
breaking (think \looseness) to gain further flexibility.

This time around I like to demonstrate new results where Alice goes “float-
ing” and all the beautiful pictures of the orginal magically appear in their ap-
propriate place.

So slowly the evolves into a usable solution. Sit back, relax, and enjoy.