Improving TEX's Typeset Layout

Hàn Thế Thành
Faculty of Informatics
Masaryk University
Brno, Czech Republic

Abstract

This paper describes an attempt to improve TEX's typeset layout in PdfTEX, based on the adjustment of interword spacing after the paragraphs have been broken into lines. Instead of changing only the interword spacing in order to justify text lines, we also slightly expand the fonts on the line as well in order to minimise excessive stretching of the interword spaces. This font expansion is implemented using horizontal scaling in PDF. When such expansion is used conservatively, and by employing appropriate settings for TEX's line-breaking and spacing parameters, this method can improve the appearance of TEX's typeset layout.

Motivation

There exist many techniques which can be used to produce high quality typeset layout. Most of these are already implemented in TEX, such as ligatures, kerning, automatic hyphenation, and very importantly the algorithm for breaking paragraphs into lines in an optimal way, generally known as "optimum fit".

However, it is still a very difficult task to obtain a uniform level of grayness of the typeset layout, even with the help of these techniques. The primary reason is that it is not possible to ensure that all the interword spaces in different lines are the same. The "optimum fit" algorithm can break the paragraph into lines in the best way, but the amount of interword space depends strongly on many other parameters, such as the paragraph width, the tolerance of glue stretching/shrinking, the amount of interword glue, etc. Considerable effort is often required in order to adjust these parameters to achieve the appropriate break points and to reduce the contrast between the interword spaces in lines. The purpose of our experiment is an attempt to perform this task better by stretching or shrinking the fonts used in each line within reasonable limits. The idea is not really new, as it represents a quite common technique using electronic font scaling in order to expand text lines that do not fit the paragraph width. However this technique is also often regarded as a bad thing, since it is frequently (ab)used in order to rescue "impossible" cases, which often leads to overdoing the scaling and produces really ugly results. In our approach, we try to use this technique in a rather different way: instead of using font scaling to improve only some "really bad" lines, we try instead to produce a "relatively good" paragraph, which does not contain any lines where the interword spacing is too bad. Then we apply font scaling to each line to reduce the difference between the interword spaces in lines. The limit of font scaling must, of course, be strictly controlled: in fact, the sum of the spaces between the words on a line is often very small in comparison to the sum of the character widths on the same line, so very slightly expanding the fonts may help considerably in improving the interword spacing.

This idea can easily be integrated with TEX because of the biggest strength of TEX – the "optimum fit" algorithm which is implemented in a very flexible manner, in order to handle restrictions on many various parameters in an optimal way. In particular, we perform the implementation in PdfTEX, where the font expansion is currently carried out by horizontal scaling in PDF as a first attempt. Other approaches may be attempted in the future as time allows.

Implementation

PdfTEX is based on the original source of TEX, and employs the changefile mechanism which allows easy access to TEX's internal data structures and simple modification of the relevant program code. Generating PDF output directly from TEX is also an advantage for our task, as we can control the spacing much better than would have been the case had we attempted it via DVI. The process of adjusting interword spacing is as follows:

It was terribly cold and nearly dark on the last evening of the old year, and the snow was falling fast. In the cold and the darkness, a poor little girl, with bare head and naked feet, roamed through the streets. It is true she had on a pair of slippers when she left home, but they were not of much use. They were very large, so large, indeed, that they had belonged to her mother, and the poor little creature had lost them in running across the street to avoid two carriages that were rolling along at a terrible rate. One of the slippers she could not find, and a boy seized upon the other and ran away with it, saying that he could use it as a cradle, when he had children of his own. So the little girl went on with her little naked feet, which were quite red and blue with the cold. In an old apron she carried a number of matches, and had a bundle of them in her hands. No one had bought anything of her the whole day, nor had any one given her even a penny. Shivering with cold and hunger, she crept along; poor little child, she looked the picture of misery. The snowflakes fell on her long, fair hair, which hung in curls on her shoulders, but she regarded them not.

It was terribly cold and nearly dark on the last evening of the old year, and the snow was falling fast. In the cold and the darkness, a poor little girl, with bare head and naked feet, roamed through the streets. It is true she had on a pair of slippers when she left home, but they were not of much use. They were very large, so large, indeed, that they had belonged to her mother, and the poor little creature had lost them in running across the street to avoid two carriages that were rolling along at a terrible rate. One of the slippers she could not find, and a boy seized upon the other and ran away with it, saying that he could use it as a cradle, when he had children of his own. So the little girl went on with her little naked feet, which were quite red and blue with the cold. In an old apron she carried a number of matches, and had a bundle of them in her hands. No one had bought anything of her the whole day, nor had any one given her even a penny. Shivering with cold and hunger, she crept along; poor little child, she looked the picture of misery. The snowflakes fell on her long, fair hair, which hung in curls on her shoulders, but she regarded them not.

Figure 1: Parameters used in this experiment: `\pdfadjustlimit` = 50, `\tolerance` = 200, `\spaceskip` = `\fontdimen2\font` plus 2`\fontdimen3\font`

Hàn Thế Thành

It was terribly cold and nearly dark on the last evening of the old year, and the snow was falling fast. In the cold and the darkness, a poor little girl, with bare head and naked feet, roamed through the streets. It is true she had on a pair of slippers when she left home, but they were not of much use. They were very large, so large, indeed, that they had belonged to her mother, and the poor little creature had lost them in running across the street to avoid two carriages that were rolling along at a terrible rate. One of the slippers she could not find, and a boy seized upon the other and ran away with it, saying that he could use it as a cradle, when he had children of his own. So the little girl went on with her little naked feet, which were quite red and blue with the cold. In an old apron she carried a number of matches, and had a bundle of them in her hands. No one had bought anything of her the whole day, nor had any one given her even a penny. Shivering with cold and hunger, she crept along; poor little child, she looked the picture of misery. The snowflakes fell on her long, fair hair, which hung in curls on her shoulders, but she regarded them not.

It was terribly cold and nearly dark on the last evening of the old year, and the snow was falling fast. In the cold and the darkness, a poor little girl, with bare head and naked feet, roamed through the streets. It is true she had on a pair of slippers when she left home, but they were not of much use. They were very large, so large, indeed, that they had belonged to her mother, and the poor little creature had lost them in running across the street to avoid two carriages that were rolling along at a terrible rate. One of the slippers she could not find, and a boy seized upon the other and ran away with it, saying that he could use it as a cradle, when he had children of his own. So the little girl went on with her little naked feet, which were quite red and blue with the cold. In an old apron she carried a number of matches, and had a bundle of them in her hands. No one had bought anything of her the whole day, nor had any one given her even a penny. Shivering with cold and hunger, she crept along; poor little child, she looked the picture of misery. The snowflakes fell on her long, fair hair, which hung in curls on her shoulders, but she regarded them not.

Figure 2: Parameters used in this experiment: `\pdfadjustlimit` = 50, `\tolerance` = 500, `\spaceskip` = `\fontdimen2\font` plus`\fontdimen3\font`

It was terribly cold and nearly dark on the last evening of the old year, and the snow was falling fast. In the cold and the darkness, a poor little girl, with bare head and naked feet, roamed through the streets. It is true she had on a pair of slippers when she left home, but they were not of much use. They were very large, so large, indeed, that they had belonged to her mother, and the poor little creature had lost them in running across the street to avoid two carriages that were rolling along at a terrible rate. One of the slippers she could not find, and a boy seized upon the other and ran away with it, saying that he could use it as a cradle, when he had children of his own. So the little girl went on with her little naked feet, which were quite red and blue with the cold. In an old apron she carried a number of matches, and had a bundle of them in her hands. No one had bought anything of her the whole day, nor had any one given her even a penny. Shivering with cold and hunger, she crept along; poor little child, she looked the picture of misery. The snowflakes fell on her long, fair hair, which hung in curls on her shoulders, but she regarded them not.

It was terribly cold and nearly dark on the last evening of the old year, and the snow was falling fast. In the cold and the darkness, a poor little girl, with bare head and naked feet, roamed through the streets. It is true she had on a pair of slippers when she left home, but they were not of much use. They were very large, so large, indeed, that they had belonged to her mother, and the poor little creature had lost them in running across the street to avoid two carriages that were rolling along at a terrible rate. One of the slippers she could not find, and a boy seized upon the other and ran away with it, saying that he could use it as a cradle, when he had children of his own. So the little girl went on with her little naked feet, which were quite red and blue with the cold. In an old apron she carried a number of matches, and had a bundle of them in her hands. No one had bought anything of her the whole day, nor had any one given her even a penny. Shivering with cold and hunger, she crept along; poor little child, she looked the picture of misery. The snowflakes fell on her long, fair hair, which hung in curls on her shoulders, but she regarded them not.

Figure 3: Parameters used in this experiment: `\pdfadjustlimit` = 30, `\tolerance` = 500, `\spaceskip` = `\fontdimen2\font` plus`\fontdimen3\font` minus`\fontdimen4\font`

- After breaking a paragraph into lines, we mark hboxes containing text lines created in this phase as boxes that might need adjustment for interword spacing.

- During glue setting of each hbox, we check whether it has been marked in the previous step. For every marked hbox we calculate the amount of font expansion for the box, depending on the sum of the character widths contained in the box and the amount of stretching/shrinking for the box.

- Finally, when shipping out marked boxes, we expand the font using horizontal scaling in PDF.

Thus the adjustment is applied only to those boxes that have been created when breaking paragraphs into lines. Also, we adjust only those boxes that need such correction: boxes with infinite stretchability and shrinkability are not changed.

There are two new primitives controlling this additional adjustment: a positive value of an integer parameter `\pdfadjustspacing` turns the adjustment on, and the value of an integer parameter `\pdfadjustlimit` specifies the limit of font expansion in thousandths of the original font width. For example, a value 50 of `\pdfadjustlimit` means that the font expansion must not exceed 5% of the original font width.

The font scaling is performed by changing the text matrix when needed. The drawback of this approach is that it makes the size of the PDF output larger. Displaying and printing such output files also takes more time.

Experimental Results

The text in the experiments was taken from the The Little Match-Seller by Hans Christian Andersen. All tests were run using the font Utopia-Regular at 11pt.

Parameters that were used to adjust line-breaking and interword spacing are indicated for each run. The left column is typeset normally, whilst the right column is typeset with font adjustment turned on. The common setting for all tests is `\frenchspacing`, `\hsize = 2.4 in` and `\emergencystretch = 1 em`.

Testing has suggested that the difference between the maximum values for stretching and shrinking should not exceed $50-60$, otherwise the font scaling will be visible and give very ugly results. Thus `\pdfadjustlimit` should not be set to a value greater than 50 while adjusting paragraphs where all the interword spaces are either all shrunk or all stretched. In cases where interword spaces can be shrunk as well as stretched, the value of `\pdfadjustlimit` should not exceed $25-30$.

The use of Multiple Master fonts might help by allowing the above limits to be exceeded; however we do not expect too much. In some cases where TeX's standard typesetting results were awful, the adjustment did not help very much.

Conclusion

Our experiments have shown that the "optimum fit" algorithm used in TeX is much more powerful and useful than it might seem. The mechanism described above can be also used to achieve better typeset layout rather than to correct bad cases in an automatic way. Trying to adjust the interword spacing is not of much use in avoiding or reducing the need of for hand-tuning of line-breaking and spacing parameters. With some effort to establish appropriate values, it can considerably enhance the uniform level of overall greyness of the typeset results.