

Chapter 10

Words to Paragraphs

We have learned how to design individual characters of a typeface using lines and curves, and how to combine them into lines. Now we must combine the lines into paragraphs, and the paragraphs into pages. Look at the following two paragraphs from Franz Kafka's *Metamorphosis*:

One morning, when Gregor Samsa woke from troubled dreams, he found himself transformed in his bed into a horrible vermin. He lay on his armour-like back, and if he lifted his head a little he could see his brown belly, slightly domed and divided by arches into stiff sections. The bedding was hardly able to cover it and seemed ready to slide off any moment. His many legs, pitifully thin compared with the size of the rest of him, waved about helplessly as he looked.

“What’s happened to me?” he thought. It wasn’t a dream. His room, a proper human room although a little too small, lay peacefully between its four familiar walls. A collection of textile samples lay spread out on the table – Samsa was a travelling salesman – and above it there hung a picture that he had recently cut out of an illustrated magazine and housed in a nice, gilded frame. It showed a lady fitted out with a fur hat and fur boa who sat upright, raising a heavy fur muff that covered the whole of her lower arm towards the viewer.

What do we notice? The left and right hand sides of the block of text are straight – no ragged edges. This is called *full justification*. We notice that some of the lines have a hyphen at the end, in the middle of a word. Looking carefully, we see that the spacing between words is not consistent from line to line. The last line of each paragraph does not go all the way to the end; the first may be indented.

How do we build a line from a list of letters? We know that each letter in a typeface has an origin, as well as an advancement which specifies how far to move to the right after drawing a character. We know also about kerning, which tells us that certain letter combinations must appear closer together. Here is a line of text, showing the (usually invisible) boxes which help to position each character:

“What’s happened to me”, he thought.


If all our characters fortuitously added up to the correct width for a line, or we were happy to break words with hyphens anywhere, or we did not want a straight right edge, this is all we would have to do. We would draw the characters in order until we reached the end of a line, and then start on the next line, moving down the page the right amount (called the *leading* – pronounced “ledding”). Alas, the world is not that simple, and we must add space to fill out the line. This can look poor if done badly, especially when a narrow column is used, such as in a newspaper:

F u l l
 justification in
 a narrow
 column can
 make big gaps
 between words
 and letters.

Here, space has been added not only between words but between letters, to make the line fit. Generally, we like to add most of the needed space between words, rather than between individual letters. Here is a paragraph typeset to three different column widths:

One morning, when Gregor Samsa woke from troubled dreams, he found himself transformed in his bed into a horrible vermin. He lay on his armour-like back, and if he...

One morning, when Gregor Samsa woke from troubled dreams, he found himself transformed in his bed into a horrible vermin. He lay on his armour-like back, and if he lifted his head a little he could see his brown belly, slightly domed and divided by arches into stiff sections.

One morning, when Gregor Samsa woke from troubled dreams, he found himself transformed in his bed into a horrible vermin. He lay on his armour-like back, and if he lifted his head a little he could see his brown belly, slightly domed and divided by arches into stiff sections.

Notice how the result improves as the column becomes wider; fewer compromises have to be made. In fact, no hyphens at all were required in the widest case. In the narrowest column, we have refused to add extra space between the letters of the compound word “armour-like”, but chose rather to produce an underfull line in this case. This decision is a matter of taste, of course. Another option is to give up on the idea of straight left and right edges, and set the text *ragged-right*. The idea is to make no changes in the spacing of words at all, just ending a line when the next word will not fit. This also eliminates hyphenation. Here is a paragraph set first ragged right, and then fully justified:

One morning, when Gregor Samsa woke from troubled dreams, he found himself transformed in his bed into a horrible vermin. He lay on his armour-like back, and if he lifted his head a little he could see his brown belly, slightly domed and divided by arches into stiff sections.

One morning, when Gregor Samsa woke from troubled dreams, he found himself transformed in his bed into a horrible vermin. He lay on his armour-like back, and if he lifted his head a little he could see his brown belly, slightly domed and divided by arches into stiff sections.

If we decide we must hyphenate a word because we cannot stretch or shrink a line without making it too ugly, how do we choose where to break it? We could just hyphenate as soon as the line is full, irrespective of where we are in the word. In the following example, the paragraph on the left prefers hyphenation

at any point to adding or removing space between words. The paragraph on the right follows usual typesetting and hyphenation rules, preferring the adding of space to hyphenation.

One morning, when Gregor Samsa woke from troubled dreams, he found himself transformed in his bed into a horrible vermin. He lay on his armour-like back, and if he lifted his head a little he could see his brown belly, slightly domed and divided by arches into stiff sections.

One morning, when Gregor Samsa woke from troubled dreams, he found himself transformed in his bed into a horrible vermin. He lay on his armour-like back, and if he lifted his head a little he could see his brown belly, slightly domed and divided by arches into stiff sections.

These are very ugly hyphenations, however: we have “troubled”, “hims-elf”, and “brow-n”. Every word has places which are better or worse for hyphenation. We would prefer “trou-bled” and “him-self”. Ideally “brown” should not be hyphenated at all. Some words must be hyphenated differently depending on context: “record” for the noun, “re-cord” for the verb, for example. In addition, authorities on hyphenation (such as dictionaries which include hyphenation information) do not always agree: Webster has “in-depend-ent” and “tri-bune”, American Heritage has “in-de-pend-ent” and “trib-une”. There are words which should never be hyphenated. For example, there is no really good place to break “squirm”.

There are two methods for solving this problem automatically as the computer typesets the lines: a *dictionary-based* system simply stores an entire word list with the hyphenation points for each word. This ensures perfect hyphenation for known words, but does not help us at all when a new word is encountered (as it often is in scientific or technical publications, or if we need to hyphenate a proper noun, such as the name of a person or city). The alternative is a *rule-based* system, which follows a set of rules about what are typically good and bad breaks. For example “a break is always allowable after “q” if followed by a vowel” or “a hyphen is fine before -ness” or “a hyphen is good between “x” and “p” in all circumstances”. We may also have *inhibiting* rules such as “never break b-ly”. Some patterns may only apply at the beginning or end of a word, others apply anywhere. In fact, these rules can be derived automatically from a list of the correct hyphenations, and be expected to work well for other unknown words (assuming

those words are in the same language – we require a hyphenation dictionary for each language appearing in the document). For example, in the typesetting system used for this book, there are 8527 rules, and only 8 exceptional cases which must be listed explicitly:

uni-ver-sity	ma-nu-scripts
uni-ver-sit-ies	re-ci-pro-city
how-ever	through-out
ma-nu-script	some-thing

Thus far, we have assumed that decisions on hyphenation are made once we reach the end of a line and find we are about to overrun it. If we are, we alter the spacing between words, or hyphenate, or some combination of the two. And so, at most we need to re-typeset the current line. Advanced line breaking algorithms use a more complicated approach, seeking to optimise the result for a whole paragraph. (We have gone line-by-line, making the best line we can for the first line, then the second etc.) It may turn out that an awkward situation later in the paragraph is prevented by making a slightly less-than-optimal decision in an earlier line, such as squeezing in an extra word or hyphenating in a good position when not strictly required. We can assign “demerits” to certain situations (a hyphenation, too much or too little spacing between words, and so on) and optimise the outcome for the least sum of such demerits. These sorts of optimisation algorithms can be quite slow for large paragraphs, taking an amount of time equal to the square of the number of lines in the paragraph. For normal texts, this is not a problem, since we are unlikely to have more than a few tens of lines in a single paragraph.

We have now dealt with splitting a text into lines and paragraphs, but similar problems occur when it comes to fitting those paragraphs onto a page. There are two worrying situations: when the last line of a paragraph is “widowed” at the top of the next page, and when the first line of a paragraph is “orphaned” on the last line of a page. Examples of a widow and an orphan are shown on the next page. It is difficult to deal with these problems without upsetting the balance of the whole two-page spread, but it can be done by slightly increasing or decreasing line spacing on one side. Another option, of course, is to edit the text, and you may be surprised to learn how often that happens.

Further small adjustments and improvements to reduce the amount of hyphenation can be introduced using so-called *microtypography*. This involves stretching or shrinking the individual char-

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulpate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultrices et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulpate metus eu enim. Vestibulum pellentesque felis eu massa.

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae tincidunt ultrices. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus

Etiam ac leo a risus tristique nonummy. Donec dignissim tincidunt nulla. Vestibulum rhoncus molestie odio. Sed lobortis, justo et pretium lobortis, mauris turpis condimentum augue, nec ultrices nibh arcu pretium enim. Nunc purus neque, placerat id, imperdiet sed, pellentesque nec, nisl. Vestibulum imperdiet neque non sem accumsan laoreet. In hac habitasse platea dictumst. Etiam condimentum facilis libero. Suspendisse in elit quis nisl aliquam dapibus. Pellentesque auctor sapien. Sed egestas sapien nec lectus. Pellentesque vel dui vel neque bibendum viverra. Aliquam porttitor nisl nec pede. Proin mattis libero vel turpis. Donec rutrum mauris et libero. Morbi lectus porta felis. Nam lobortis, metus quis elementum commodo, nunc lectus elementum mauris, eget vulpate ligula tellus eu neque. Vivamus eu dolor.

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulpate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultrices et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat

quis tortor vitae risus porta vehicula.

Fusee mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultrices tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in, hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilis. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetur.

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula, urna sed ultrices auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur et nunc. Aliquam dolor odio, commodo pretium, ultrices non, pharetra in, velit. Integer arcu est, nonummy in, fermentum faucibus, egestas vel, odio.

Sed commodo posuere pede. Mauris ut est. Ut quis purus. Sed ac odio. Sed vehicula hendrerit sem. Duis non odio. Morbi ut dui. Sed accumsan risus eget odio. In hac habitasse platea dictumst. Pellentesque non elit. Fused sed justo eu urna porta tincidunt. Mauris felis odio, sollicitudin sed, volutpat a, ornare ac, erat. Morbi quis dolor. Donec pellentesque, erat ac sagittis semper, nunc dui lobortis purus, quis congue purus metus ultrices tellus. Proin et quam. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos luptentates. Praesent sapien turpis, fermentum vel, eleifend faucibus, vehicula eu, lacus.

Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Donec odio elit, dictum in, hendrerit sit amet, egestas sed, leo. Praesent feugiat sapien aliquet odio. Integer vitae justo. Aliquam vestibulum fringilla lorem. Sed neque lectus, consectetur at, consectetur sed, eleifend ac, lectus. Nulla facilis. Pellentesque eget lectus. Proin et metus. Sed porttitor. In hac habitasse platea dictumst. Suspendisse eu lectus. Ut mi mi, lacinia sit amet, placerat et, mollis vitae, dui. Sed ante tellus, tristique ut, iaculis in, malesuada ac, dui. Mauris nibh leo, facilisis non, adipiscing quis, ultrices a, dui.

Morbi luctus, wisi viverra faucibus pretium, nibh est placerat odio, nec congue sed, wisi enim eget quam. Quisque libero justo, consectetur a, feugiat vitae, porttitor eu, libero. Suspendisse sed mauris vitae elit sollicitudin malesuada. Maecenas ultrices eros sit amet ante. Ut venenatis velit. Maecenas

at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulpate metus eu enim. Vestibulum pellentesque felis eu massa.

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Fusee mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultrices tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in, hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilis. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetur.

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula, urna sed ultrices auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur et nunc. Aliquam dolor odio, commodo pretium, ultrices non, pharetra in, velit. Integer arcu est, nonummy in, fermentum faucibus, egestas vel, odio.

Sed commodo posuere pede. Mauris ut est. Ut quis purus. Sed ac odio. Sed vehicula hendrerit sem. Duis non odio. Morbi ut dui. Sed accumsan risus eget odio. In hac habitasse platea dictumst. Pellentesque non elit. Fused sed justo eu urna porta tincidunt. Mauris felis odio, sollicitudin sed, volutpat a, ornare ac, erat. Morbi quis dolor. Donec pellentesque, erat ac sagittis semper, nunc dui lobortis purus, quis congue purus metus ultrices tellus. Proin et quam. Class aptent taciti sociosqu ad litora torquent per

A widow (top) and orphan (bottom).

acters in a line, hoping to make the line fit without the need for hyphenation. Of course, if taken to extremes, this would remove all hyphens, but make the page unreadable! Shrinking or stretching by up to 2% seems to be hard to notice, though. Can you spot the use of microtypography in the paragraphs of this book?

Another way to improve the look of a paragraph is to allow punctuation to hang over the end of the line. For example, a comma or a hyphen should hang a little over the right hand side – this makes the block of the paragraph seem visually more straight, even though really we have made it less straight. Here is a narrow paragraph without overhanging punctuation (left), then with (middle):

<p>One morning, when Gregor Samsa woke from troubled dreams, he found himself trans- formed in his bed into a horrible vermin. He lay on his armour-like back, and if he lifted his head a little he could see his brown belly, slightly domed and divided...</p>	<p>One morning, when Gregor Samsa woke from troubled dreams, he found himself trans- formed in his bed into a horrible vermin. He lay on his armour-like back, and if he lifted his head a little he could see his brown belly, slightly domed and divided...</p>	<p>One morning, when Gregor Samsa woke from troubled dreams, he found himself trans- formed in his bed into a horrible vermin. He lay on his armour-like back, and if he lifted his head a little he could see his brown belly, slightly domed and divided...</p>
---	---	---

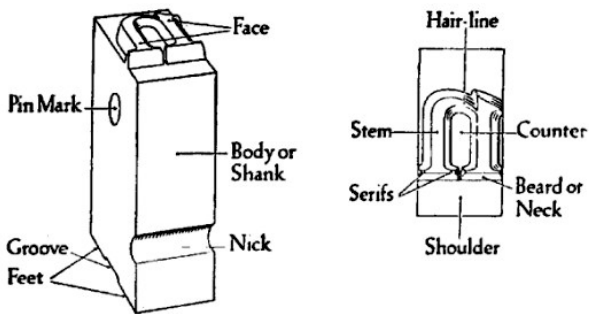
The vertical line (far right) highlights the overhanging hyphens and commas used to keep the right hand margin visually straight. A further distracting visual problem in paragraphs is that of *rivers*. These are the vertical lines of white space which occur when spaces on successive lines are in just the wrong place:

interdum sem, dapibus sempe
et augue. Quisque cursus nul
auris malesuada mollis. Nullæ
; arcu, egestas ac, fermentum c
s, sem in pretium fermentum,
;equat augue urna in wisi. Qui
uis, lacinia in, est. Fusce facilis

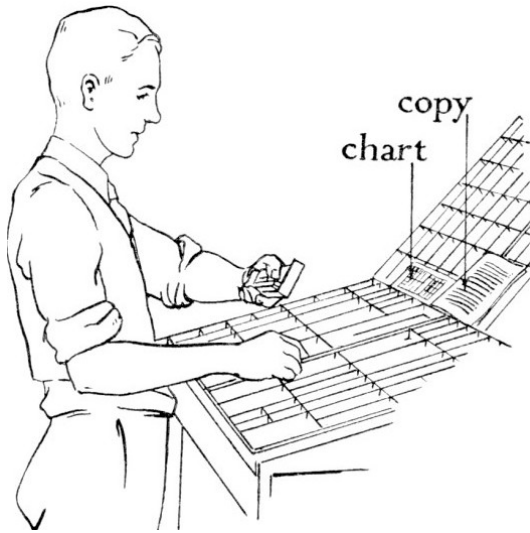
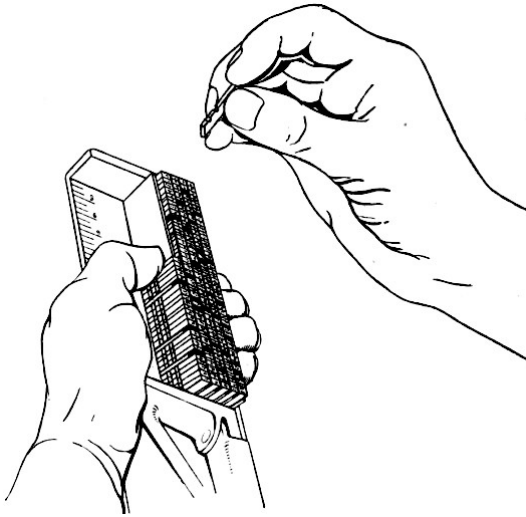
interdum sem, dapibus sempe
 et augue. Quisque cursus nul
 lauris malesuada mollis. Nulla
 ; arcu, egestas ac, fermentum c
 s, sem in pretium fermentum,
 sequat augue urna in wisi. Qui
 quis, lacinia in, est. Fusce facilis

We have shown the river with a line. Notice that the word “fermen-
 tum” appearing in almost the same place on two successive lines is
 also distracting. The problem is difficult to deal with automatically,
 and the text may have to be edited to fix it. The microtypographical
 techniques discussed above can help a little – since there are fewer
 widened spaces between words, the rivers will be narrower and
 less noticeable.

You may wonder how type was set before computers. In much
 the same way, it turns out, but with many more manual steps and a
 lot of little pieces of metal. Here is one such piece, for the character
 “n” at a particular size, in a particular typeface:



These are picked from a tray of boxes, by hand, and placed into
 rows into a *composing stick*, each word separated by little metal
 spaces, each row spaced by a metal strip (the leading). You can
 imagine that many many copies of these little metal pieces were
 required for each typeface and size, so it was an expensive business.
 Because it will eventually be used for printing by being inked and
 stamped or rolled on paper, the type is mirrored, and hard to read,
 and one must be careful not to mix up “p” and “q”, or “b” and “d”.
 (This is one possible origin of the phrase “mind your Ps and Qs”.)
 This painstaking process is shown on the opposite page.



The finished paragraphs of type are arranged in a *galley*. This will be used to make prints of the page (or pages – two or four may be printed from one galley, then folded and cut). You can imagine how long it takes to make up the galleys for a book, and how much time is required to justify each line by inserting exactly the right spaces and hyphenating by hand. Mistakes found after test prints can be very costly to fix, since they necessitate taking apart the

galley and replacing not just a single character, but perhaps re-typesetting a whole paragraph. Here is a galley, ready for printing:



Eventually, machines were developed to automatically place the pieces of type based on what was typed on a keyboard and to automatically justify each line. Such mechanical systems were in common use until the advent of so-called phototypesetting. This involved building an image by shining light through a series of stencils onto photosensitive paper, then photographing it. Computer typesetting supplanted both in the late twentieth century.

Problems

Solutions on page 166.

Identify good hyphenation points in the following words:

1. hyphenation
2. fundraising
3. arithmetic (noun)
4. arithmetic (adjective)
5. demonstration
6. demonstrative
7. genuine
8. mountainous