

8pen research

Whenever we have an idea, or want to take a note, we reach out for pen and paper. It is often the fastest, and easiest way take a note. We enjoy handwriting, because it is natural. One may ask:

Main Problem: Is there a natural way to write on small devices?

In other words, is it possible to maintain the experience of handwriting, while using digital devices? We should hope so. While the possibilities are limitless, we wish to specify:

More specific Problem: Can we devise a method of writing on small touch-enabled devices that mimics handwriting?

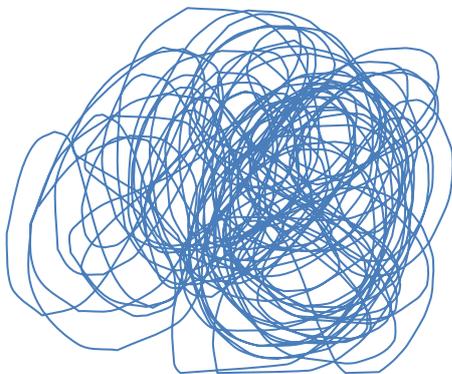
It is important to note that “mimics” here only refers to the essential features of handwriting as a human experience, namely its ease of use, its natural application, its flexibility and its propensity to be enjoyed as a creative activity, and less to its specific gestures.

Let us take a look at a sample of a handwritten note:

Enjoy writing.

It is a continuous line, drawn in a steady movement from left to right. The first problem that you encounter when you try to “imitate” that naively on a small device is that there just isn’t enough space.

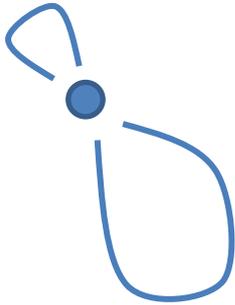
Simple Observation: When asked to draw a continuous line in steady motion on a small square canvas, most people revert to circular motions and changes of directions resembling the shape of an 8.



If one carries out this motion for a longer period of time, and steps back from the resulting drawing on the canvas, we can see a dot emerging in the centre, a region where the line passes through most frequently from all directions, and a ring of circular line segments surrounding this dot.

8pen research

Proposition: The natural continuous motion constrained to a small square consists conceptually of a number of *loops* glued together at a central dot.



Signature Problem: What is the easiest way to put a signature on the loops?

The problem simply aims at an easy classification of all possible loops. How many are there, and what is a sensible way to distinguish them?

Simple Resolution: The loops differ in (i) their direction of motion, (ii) how far they go around the dot, and (iii) their orientation (if we think of them rotated around the dot). With these 3 properties in place, the number of all possible loops is already hugely reduced, but we shall work with yet a smaller class of loops, by discretizing our choices in (ii), (iii).

The idea of sectors. Divide the area around the dot into a number n of sectors of the circle, like a cake into pieces. We then distinguish only between loops that took $1/n, 2/n, \dots$ of the full turn, and start and end in one of the sectors. In essence, this introduces an equivalence of loops, based on the numbers of sectors they have passed through (in any given direction), before returning to (and after leaving) the dot from (to) any of the sectors. The obvious question here is, what number should n be? We can answer that, in view of the following:

Basic Principle: To each loop we associate a letter.

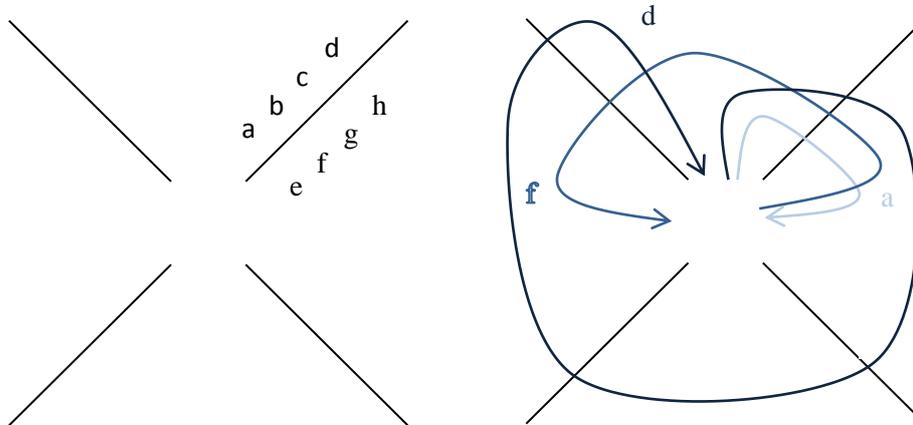
Since a typical alphabet has around 30 letters, a system with n chosen to be less or equal to 3 would lead to the assignment of letters to loops with a length exceeding the full turn, while a system with n greater or equal to 5 would leave many loops unused. With n equals 4, we have 32 loops to play with.

Assignment Problem: What letters should be associated to which loops?

This is not easy. As we shall see this is a problem lying at the intersection of the fields of mathematical optimization, linguistics, and ergonomics. Linguistics provides us with the knowledge of the structure and usage of language and information about typical arrangements of letters, while ergonomic considerations enter in deciding how this should be reflected in the movements of the hand. More precisely, it is the language which tells us which sequences of letters occur how frequently and in which context, and given an assignment of letters to loops, it is then a matter of ergonomics to assess in which sense this assignment is "optimal". With a notion of optimality and a collection of assignments at hand, this becomes a problem in mathematical optimization.

8pen research

Layouts. We can represent a particular assignment of letters to loops on a 4-sector canvas as follows:



The canvas is divided by four edges into four sectors. As is illustrated for one of the edges, the letters lined up along one side of it indicate an assignment of these letters to the loops starting in the sector to that side of the edge, and leaving the sector on the side of their alignment, in ascending order of the sectors these loops pass through, the innermost indicating a return to the centre straight after having passed the edge, the outermost being associated with all loops that take a full turn in that direction.

Lemma: There are as many layouts (as many assignments of letters to loops in the present “Simple Resolution”) as there are permutations of the alphabet.

Unless you are an astronomer, you may amuse yourself by calculating this number that is likely to be bigger than any number you have seen before. Even if we did have a simple way of assessing whether a layout is “good”, no computer could go through all possible layouts in reasonable time.

A notion of optimality. An easy way to tell if a layout is “good” or not is to parse a text through it, and to calculate its “length,” where each letter contributes with the number of sectors passed through in order to write it. The resulting number is somewhere in the interval of one to four times the actual length of the text, and the layouts can be ranked accordingly.

This is the simplest quantitative notion of optimality we can think of, but it already raises all the fundamental questions related to approaches in that direction:

Sample dependence: To what extent does the resulting ranking depend on the sample of text that was used? And what text should be used, what samples should a layout be optimized for?

Open research

Definitions of length: There are many ways of assigning a “length” to a text in a given layout, the above being only the simplest. Similarly, what is the dependence of the result on the notion of “length,” and which are equivalent? More importantly however, what is the definition of “length,” that yields the most *ergonomic* layout?

Landscape of layouts: With a chosen text sample, and an agreed notion of length, is the optimal layout unique? Imagine looking out into a landscape of layouts, each point on the map being a layout at an altitude corresponding to its length, is that landscape hilly, with many valleys, or is it more like a volcano, with just a single crater?

In our approach we have optimized mainly for *syllables*, largely inspired by the Korean alphabet, and have moved on from “length” to an appropriate notion of “*energy*”. We believe that the landscape is indeed hilly, and will give good reasons why we have settled on a *certain* valley, which may as well be the crater.

Global Statistics, and a priori selection of layouts. We can think of the letters to be arranged in shells around the centre, similar to electrons around a nucleus in the semi-classical model of the atom. Indeed, each of the 4 shells holds 8 letters, and we may arrive at a layout by filling up these shells, one after another. The innermost shell should contain the most frequently used letters, while the outer shells should contain the less frequently used, for the inner shells are associated with the smallest, and thus fastest loops. This leads us to ask a simple question in “quantitative” comparative linguistics:

Letter Frequency: What are the most frequently used letters in any given language? More specifically, is there a division of the latin alphabet in groups of at most 8 letters reflecting their frequency of occurrence that is compatible with the “common usage” of any language using the latin alphabet?

Unfortunately, the latter can’t be strictly true, but still in this respect the languages based on the latin alphabet offer a reasonable compromise. Let us first take a look at English, and the deviations in for example French, German, and Spanish:

	English	French	German	Spanish
Shell 1	e, t, a, o, i, h, n, s	u, r (2)	r, d (2)	r, l (2)
Shell 2	r, d, l, u, m, w, y, f	o (1); c, p, v, q (3)	h, o (1); c, g, b (3)	t (1); c, p, b, v (3)
Shell 3	c, g, p, b, v, k, x, q	h (1); f, y (2); j, z (4)	f, w, y (2); z, j (4)	h (1); f, y (2); j, z (4)
Shell 4	j, z	w (2); k (3)	q, x (3)	w (2); k (3)

Ideally, if the languages French, German, and Spanish were “compatible” with English in the above sense, then their respective columns would be blank; (for each shell only those letters are displayed that are not already found in the corresponding reference shell, and it is displayed in brackets from which shell they “come from” in the English “atom”). However, it is mostly due to a few exceptions of letters (here notably ‘h’, ‘y’, ‘w’) that are “knocked out” from the inner shells, which then causes the letters in the outer shell to “fall” further inside (like ‘r’, ‘c’, ‘z’ etc); (conversely it happens that a specific letter “jumps” into an inner shell, and “pushes” the other letters outwards, like ‘q’ in French).

Open research

The entire purpose of this point of view is that it radically reduces the number of possible layouts, by restricting the permutations of the alphabet to those that respect a chosen structure of the shells. This restriction may be done incrementally, first assigning only a few letters to certain shells (a suitable selection that maximizes the compatibility among a set of languages in the above sense) while keeping all the remaining letters loose.

While the arrangement of the letters into shells with the help of “global statistics” makes the problem tractable, the key to its success really lies in the placement of the letters within a given shell.

Hangul, and the idea of syllables. “There is something gratifying about finishing a word in a good way,” a writer might say reflecting about his own handwriting. Or have you ever observed someone drawing her signature? We like to speed up towards the end, often the last few letters of a handwritten word resemble very much a single stroke. We finish a word in our mind, before our hand does.

Intuitive Syllable Assumption: Human handwriting is conducted in such a way that we break up words into syllables, and settle into a rhythm where little pauses are taken before suffixes or after prefixes. In other words, we are guided by the idea that the morphology of a word manifests itself in the way it is hand-written.

A lucid illustration of this concept is found in the Korean way of writing, using Hangul. Here, typically a handful of letters are grouped together, and assembled in a structured way to form syllables, which then in turn form words.

Example	I like Korean
Korean letter sequence	ㅎ ㄴ ㄱ ㅏ ㅓ ㅗ ㅛ ㅜ ㅠ ㅇ ㅈ ㅊ ㅎ ㅇ ㅍ ㅇ ㅂ
Hangul	한국어 좋아해요

We are lead to conclude that while the overall frequency of occurrence of a letter in a given language should determine to what shell it belongs, it is a matter of its appearances in the most frequent affixes that governs its placement within a shell.

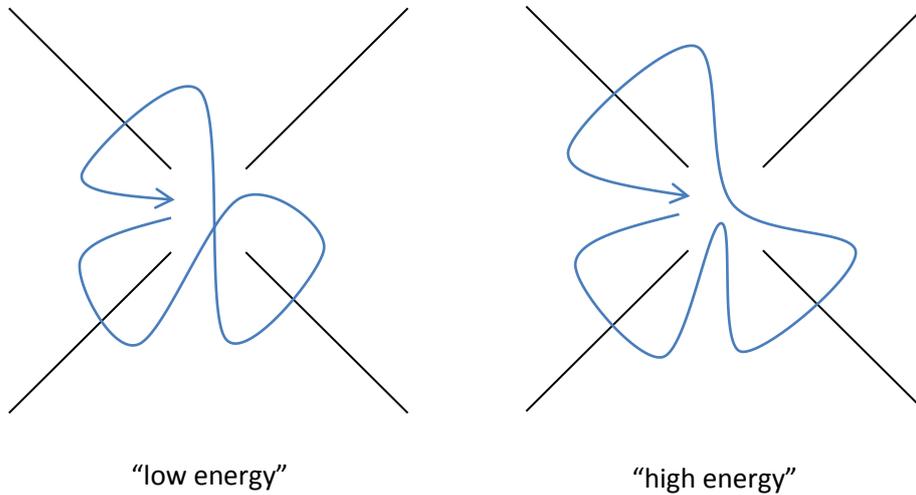
Examples of Affixes	Prefixes	Suffixes
English	pre, in, re, un	ed, able, al, ly, an, ful, ing, ish, ism, ity, ize, less, ly, ness
French	co, de, en, im, in, pre, re	able, age, ais, ance, en, er, ette, eau, eur, eux, ien, ier, ique, iste, ite, ieme, ment, ois, on
German	ab, aus, be, un, ver	bar, er, heit, ig, in, keit, lich, ung
Spanish	des, in	al, ar, dad, eria, ista, mente, on

What we are aiming for is primarily *fluidity*. The affixes of a language should become one single gesture, and any given phrase should be written smoothly. This is what makes writing fast, and enjoyable.

Ergonomics, and a notion of energy. Each letter is associated to a loop; for an ending like “ing” to become a single gesture means that the loops of ‘i’, ‘n’, and ‘g’ *fit together*. While length of this

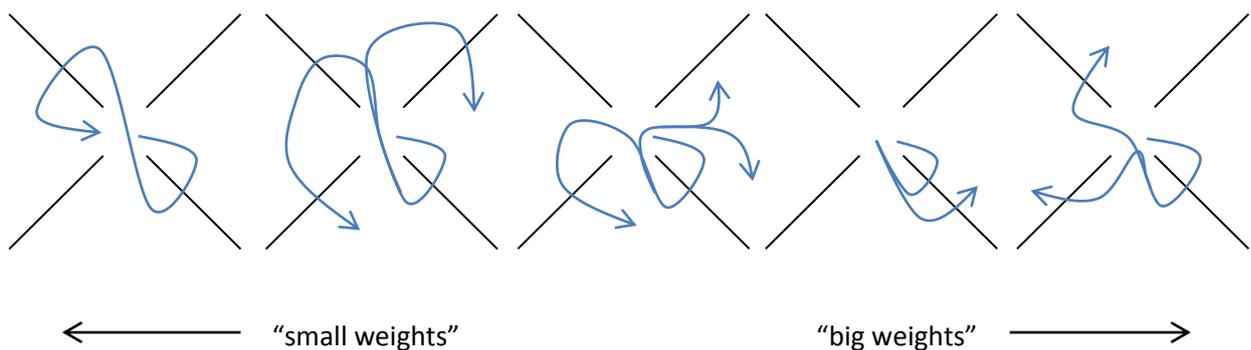
8pen research

syllable may be the total number of sectors passed through by the corresponding loops, what really matters is this syllable's "energy", of an appropriately defined type, which is less concerned with a single loop, but rather with the succession of loops.



Definition: The energy of a word is the sum of the energies of its syllables. Moreover, the energy of a syllable is the total length of the loops it consists of (the total number of sectors passed through by these loops) times a "weight", a factor that evaluates the "ergonomic smoothness" of the succession of the loops.

Note that this is to the effect that a syllable may be "long", but if it occurs as a smooth gesture then its energy can be very "low". More specifically, we propose:



We are thus constructing a candidate for the solution of the "Assignment Problem" as a minimizer of a suitably chosen energy function.

Definition: Associated to each collection of words \mathcal{C} there is an *energy function* $E_{\mathcal{C}}: \mathcal{L} \rightarrow \mathbb{R}$, defined on the set of all possible layouts \mathcal{L} , and taking values in the real numbers \mathbb{R} , by setting $E_{\mathcal{C}}(L)$ to be

8pen research

the sum of the energies of all words in \mathcal{C} on the layout $L \in \mathcal{L}$ in the above sense. For any subset $\mathcal{S} \subset \mathcal{L}$, we call $M \in \mathcal{S}$ a *minimizer* on \mathcal{S} for the collection \mathcal{C} if $E_{\mathcal{C}}(M) = \min_{L \in \mathcal{S}} E_{\mathcal{C}}(L)$. Moreover, let us refer to the “good” layouts to mean those that are in $\mathcal{G}_{\mathcal{C}}^{\mathcal{S}} = \{L \in \mathcal{S} : E_{\mathcal{C}}(L) \leq C \times E_{\mathcal{C}}(M)\}$ for an appropriately chosen constant $C > 1$.

Main Result: The presently used layout L^* for the 8pen is an element of $\mathcal{G}_{\mathcal{A}}^{\mathcal{S}} \cap \mathcal{G}_{\mathcal{T}}^{\mathcal{S}} \cap \mathcal{G}_{\mathcal{P}}^{\mathcal{S}}$, where \mathcal{A} is a collection of affixes taken from various languages using the latin alphabet, \mathcal{T} a longer text representative for these languages, and \mathcal{P} a collection of common phrases. Here $\mathcal{S} \subset \mathcal{L}$ is the previously discussed subset of a priori selected layouts (of letters arranged into shells) by virtue of global statistics.

This outlines the line of thought that ultimately leads to the 8pen. However, at each point in the argument alternative solutions and ideas may be conceived, which could lead to altogether very different methods of writing, which address our initial Main Problem. Even in this approach, many questions remain open, among which we would like to mention:

Questions:

Signature Problem. Is there a more general resolution of the Signature Problem? In particular, is there a way to put a signature on loops in the absence of a distinguished central dot?

Structural linguistics. In the Korean layout, which will be discussed in a separate note, one observes a separation of the letters into vowels “to the right” and consonants “to the left”. This is not surprising in view of the ergonomic choices we have made which favour motions in the shape of an 8, and the structure of the Korean syllables which dictates an alternation of the vowels and the consonants. However, it is an example how the linguistic structure manifests itself in the layout, and the question arises which structural aspects of language should influence the conception of a layout from the beginning, and what in turn can be learned about linguistics from an optimal layout.

Genetic algorithms. One may shift the view from a “landscape” of layouts to a “population” of layouts. It is then conceivable to start out with a number of “good” albeit very different layouts, and “to cross” these layouts in the hope to find even “better” layouts in the “populations of the following generations.” Is there a natural way to pair 2 layouts, such that the “child” maintains features of its “parents”, or with an appropriate notion of optimality is at least as “good”? Since the virtues of a “good” layout are very interwoven, it is not expected that this can be done in a straight forward way; however, if there was a way to pair 2 layouts, then one could study a far wider “population” of layouts, lifting the restriction of the assignment of letters into shells.