

University of Groningen

Patterns and symbols

Schomaker, Lambertus

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version

Publisher's PDF, also known as Version of record

Publication date:

2002

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Schomaker, L. (2002). *Patterns and symbols: a world through the eye of the machine*. s.n.

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Patterns and symbols:
A world through the eye of the machine *

Lambert Schomaker
Inaugural speech,

presented at the occasion
of accepting the
office of professor
in Artificial Intelligence
at the Rijksuniversiteit Groningen
Groningen, 10 December 2002

(a coarse translation of the original text in Dutch)

Ladies & gentlemen,

at this moment in time, I realize clearly that the name of my chair, i.e., *Artificial Intelligence*, may contain an intrinsic contradiction. Indeed, after fifty years of research, it is still not clear whether humanity will be able to construct an artificial system which can perceive our world and understand it, in order to display intelligent behaviour consequently.

However, I do not worry about this possible shortcoming. Indeed it is a well-known observation that the sciences which are most remotely located from a possible practical utility will enjoy the highest esteem. Furthermore, the challenge which is presented to me would be much less stimulating if autonomous intelligent machines were already leaving the industrial conveyor belts at this very moment, with an elegant jump in order to assist us, humans.

It was the advent of the computing engine which lead to the question which keeps my research field awake: “can intelligence be constructed?” The attempts which have been undertaken to find answers have generated a huge amount of knowledge, by now. I will try today to share a small part of this knowledge with you, in the form of insights which I have developed during my own scientific journey.

However, before concentrating on the actual content of this essay, I would like to commence with a more philosophically or epistemologically oriented introduction, in order to indicate the general framework within which the later part of this presentation can be understood.

The word “science” is used frivolously as a receptacle for a wide spectrum of human efforts. As is often the case with abstract concepts, the term “science” will often elicit an ‘Aha-erlebnis’ [superficial recognition] in the brain of the listener, which is however devoid of a thorough understanding.

In my opinion it is sensible to make a distinction between three forms of ‘-ship’: *thinkingship*, *knowingship* and *skillship* *. The neologisms may sound strange to your ears but these terms demarcate the meaning which I would like to elucidate here.

Thinkingship [Denkenschap] concerns the development of systems of concepts. Ideas are developed, albeit slowly, here and there supported by experimental research. Thinkingship is a rather enjoyable pastime where one is only marginally bothered by earthly trivialities. The thinker strives for a simplification of the interpretation of the stream of information in which we are immersed during our interactions with the world. However, empiricism only plays a limited role in thinkingship. The thinker mainly aims at obtaining aestheticism in thought. If one is lucky, the thinker uses Occam’s razor, otherwise the stream of words would be massive. Indeed, in thinkingship, a cunning use is made of the word, i.e., the niches, the shortcomings and seductions that natural language may offer to push and pull the listener in a desired direction. Unfortunately, in thinkingship, mathematics is missing as a tool and as a lingua franca. Nevertheless, thinkingship has the richest tradition and is highly esteemed within the scientific community.

Real *Knowingship* [wetenschap] involves *knowledge of nature*, incorporating the two unavoidable: *logic* and *mathematics*. Actual “knowing” is however a scarce good. In some research areas good progress is made: One proceeds in small steps, with a good amount of annual publications which are of limited scope, each in itself. In a step-wise fashion, our world is enriched with new, mostly small but interesting facts from nature. Obviously there exists a risk that all these small facts are only marginally interrelated, but still it is here, of course I am hinting at biology, where the largest strides are made. In neighbouring research areas however, where one hopes to participate in science, many facts are diddled from nature, too, but as long as these facts are contrary to a pursued highly aesthetic Idea, they are not written down. The latter would be senseless anyway because the envisaged stages only allows for grandiose conceptions to be presented to their audiences. Thus, pursued science (knowingship) commonly suffers from an abundance of thinkingship. Pro forma one tests the cold water of reality with the big toe, but not too long and not too deep, because thinkingship is a more pleasurable and drier pastime than is real knowingship.

Finally, and thirdly, there is *Skillship* [kunnenschap], a young branch. Here we encounter the builders. The builders are rarely motivated by grandiosity of ideas, nor are they satisfied to ‘know that’. The builder

*Here, the author plays with the Dutch word for science, ‘wetenschap’ ending with ‘-schap’ which is etymologically related to the English postfix ‘-ship’: *denkenschap*, *wetenschap* en *kunnenschap*.

obtains satisfaction from another type of victory on nature: A cure, a well-designed construction, a humming machine or a working algorithm delivers the gratification for all work. Here, the focus is “know how”. Due to the fact that a builder concentrates all efforts on the struggle with matter, it appears that time and motivation to submerge in argumentation at the level of *thinkingship*, is often lacking. The builder enjoys victory over nature, per se, and the approval of colleagues. Today, the builders have learned to transform their “know how” into “know that”, by means of language and mathematics. With suspicion, however, the builder looks upon *thinkingship*. Indeed, through continuous contact with the resilient reality, a builder knows that striving for an esthetically pleasing system of concepts cannot be the only guiding principle to understand and control the world.

In artificial intelligence, we find all three of them: The Thinker, the Knower and the Builder, often combined in one and the same person. This is not strange in itself. Of sir Isaac Newton it is known that he computed ballistic cannonball trajectories in the morning, while spending the afternoon, inspired by such practical work, thinking about gravity. Nevertheless, there exists a problem. For a long time, natural philosophers (scientists) were not admitted to partake in *thinkingship*: Anthonie van Leeuwenhoek and Christiaan Huygens were not allowed to enter the Academie, Teyler’s Genootschap in Haarlem, easily. Their work was too earthly, as they were not regular philosophers or theologians. The builders (engineers) and the medical practitioners, both groups consisting of members of the third caste, even had to wait until the previous century before they were allowed to join the social order of academics. Why this is relevant, you wonder?

Until today, the aforementioned classification seems to exist within the order of academics, with a corresponding pecking order in which the builders unfortunately reside at the bottom. A research area such as Artificial Intelligence, and even the field of Computer Science at large, is deemed as suspect by some in academia. Indeed, both computer science and artificial intelligence have not taken nature, but a *human artifact*, i.e., the computer, as a source of inspiration for their research: “t is all trendy hype with a short life span” [such skepticism is expressed in old fields such as physics and psychology].

However, one can easily relativise this problem: It was not the falling apple, but the cannonball which constituted a major inspiration to Newton (Figure 1).

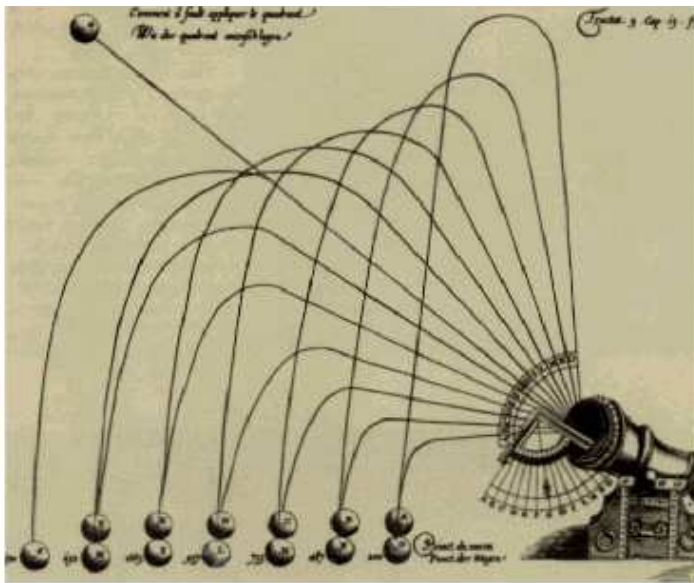


Figure 1: Here we see Newton, looking in a reprobratory manner at an overly conceptual, pre-Newtonian, i.e., non-parabolic rendering of a cannonball trajectory. Technology is an excellent inspirator for science.

The cannon, just like the computer, constitutes a technical artifact and not a given natural item. During the industrial revolution of the nineteenth century, the exploding steam engines necessitated a development of basic knowledge, leading to the respectable research field of thermodynamics. In a similar vein, the advent of the computer provides more than a mere practical utility. The “Electronic Brain” evoked a large number of challenging scientific questions. Also in this case, the presence of a man-made engine leads to the insight that there is ignorance. What are the fundamental possibilities and limitations of reasoning by means of an automaton? The attractiveness of a perfect reasoning mechanism remains to exist until today. The presence of the computer forces us to a reconsideration of the position of the human within the set of cognitive systems. The builder plays an important role in this aspect: *Understanding by Building* is an essential part of real science, with a role of ever-increasing importance, not only within computer science and artificial intelligence. Through our goal of working models of perceiving and reasoning systems we are confronted with the strong and weak sides of natural intelligence and cognition. Taking all of this in consideration, there is sufficient raw material for a thriving evolution of the interdisciplinary research field of artificial intelligence in the coming years. In Groningen, this flowering development has started about ten years ago through the enthusiastic work of an interdisciplinary team with roots in psychology, cognitive science, informatics, logic, biophysics and linguistics, under the common denominator *Cognitive Engineering*

Incommensurabilia

After this brief pleading for the existence rights of our research area, I would like to focus on the actual content of this inaugural address.

The largest fundamental problem in the development of intelligent systems which can perceive, reason and act, is the chose computational paradigm. There are two perspectives on information processing, each

entailing a rich arsenal of methods. Unfortunately, we have to position both perspectives under the class of incommensurabilia, i.e., the incommensurables.

The first method for modeling is logic, which experienced a tremendous development as an applied method due to the fact that the formalisms of Boole (1848) could be materialised. Initially, this was realized by means of relays, later by means of thermionic valves and ultimately with transistors. Boole himself stated in 1848 that he had invented:

“... a new and peculiar form of mathematics to the expression of the *operations of the mind in reasoning* ...”

One hundred year later, one courageously commences to exploit the power of logic for solving many problems in information processing by means of the new invention, the computer. That seems to work reasonably well. This even works so good indeed that optimistic predictions are being made concerning chess-playing, perceiving and language-translating computers which would populate this planet around the year 2000. Computer chess has become a great success indeed. Although there appears an occasional human master player who is able to win, the large majority of humans has lost the competition already since the last decade. Scientifically, the issue has been settled. Indeed, in the case of new medicine, one does not expect that it really heals *all* patients in order to pronounce it as effective. However, in other research areas than computer chess, an important evolution took place in the background.

The use of logic was not everywhere as successful as it was in the implementation of basic calculation and search processes by the computer. Boole had stated it so clearly: *logic is used for reasoning*. The availability of the computer, however, was deemed useful for a very wide spectrum of information-processing functions. In the initial years, logic was directly applied to the lowest processing levels: For the implementation of a visual perceptive system, the binary value of an individual image element was considered to constitute a logical proposition for a reasoning system. Such an approach would be inconceivable, today. It became quickly clear that a logic-based approach to perception entailed huge problems.

First, there is always noise in the perception of an image of the external world. There is noise in the sensor, there are vibrations, there is movement. Second, one needs powerful transforms of visual and auditory signals in order to achieve robust representations which can be used as raw material for a reasoning engine. To some researchers it became clear that the expressive power of logic failed here. Around 1969, fermentation took place in the world of cybernetics and in the beginning of the nineteen seventies the unavoidable took place: A separation into two research fields. In 1969 the first International Joint Conference on Artificial Intelligence was held. Four years later already, the renegades separated themselves from the main stream of logic resulting in the first International Joint Conference on Pattern Recognition, in 1973. Until today there exists this rather strict separation between research in machine-based reasoning and research in machine-based perception.

There existed several reasons of sociological and epistemological nature for the split between Artificial Intelligence and Pattern Recognition but most important is this question: “does it suffice to use logic as the basis for the implementation of an intelligent system or does one need other tools, from statistics, geometry, linear algebra and signal-processing theory?”. The pattern-recognition renegades who were disappointed by the limitations in the use of logic in modeling perception were proved right in many points. However, the price which was paid by this separation of fields is very high, a point to which I will return at a later stage.

In order to illustrate what type of drama evolved here, I will present a few examples from the field of automatic script recognition. Today, the recognition of machine-printed fonts does not constitute a large problem anymore if the image consists of a fronto-parallel projection of sharply discernible text in a known font and rendering. The automatic recognition of cursive and free-style handwritten text, on the other hand, still poses a difficult research challenge.



Figure 2: The word *algebra*, written by eight writers (rows), at four occasions (columns) each.

Here we see a number of handwritten words in a cursive style. Eight writers (the rows of the matrix) visit the lab at four occasions (the columns) to write the word *algebra*. The first observation that can be made concerns the difference in writing style between the writers. Additionally, it can be seen that there exists a variability of the writing product for a single writer. How to proceed, if our goal is to develop an algorithm for the recognition of words by means of an analysis at the level of letters?

In the initial years of automatic script recognition the goal was to identify topological invariants at an abstract level, such that each word or each letter could be represented uniquely. There is a hope that the found mathematical representation has general validity such that a large set of written shapes can be uniquely reduced to a single class of abstract patterns. The search for such representations or shape descriptors can be best described as a quest for the golden grail. Most research in this area is concerned with arbitrarily chosen solutions, by the Builders among us. Several attempts, however, have been made to use theoretical guidelines for the choice of shape descriptors (features). The engineers initially consulted the literature of experimental psychology. This research field, however, appeared much more concerned with self-generated problems in odd experimental settings than with the basic question *which aspects of handwritten character shape the human reader actually attends to while reading?* Out of necessity, the builders had to come up with their own theories and solutions.

An example of an interesting theory is the work of the eminent French engineer prof. Jean-Claude Simon[†], a pioneer in the field of pattern recognition. He makes a distinction between regularities and singularities. Looking at handwriting and speech he concludes that these signals of human origin consist of two components:

Firstly, a regular base shape out of which, secondly, a number of distinctive elements emerge at specific places or moments in time. In speech, this can be observed in the distinction between the vowels, i.e., the relatively long-lasting periods in which the vocal chord produces a periodic acoustic signal, intermitted by singularities, i.e., the consonants. Similarly, in handwriting, one may make a distinction between an oscillating main axis, from which at a number of points clearly discernible singularities emerge, i.e., the ascenders, descenders, crossings and other topologically unique elements (Figure 3).

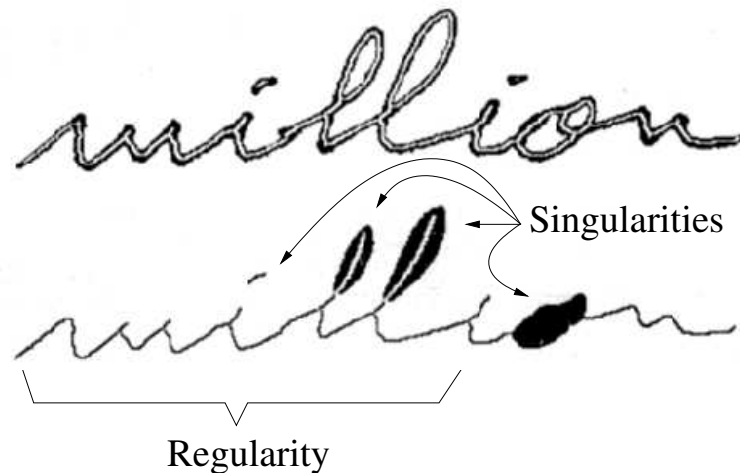


Figure 3: Singularities en regularities in a cursively handwritten word (Simon, 1991). The enclosed spaces of the letters l en o are easily detected by any visual system, thus constituting singularities. The undulating main axis of the word *million* represents a regularity which is in itself not very informative, but it presents a context for the singularities. Also a dot on a letter can be considered to constitute a singularity.

On the basis of the theory of Simon, a piece of handwritten text can be dissected into a basic shape which is not very informative in itself, and the salient singularities. By means of a symbolic description, a shape language, one may try to represent each word. A sentence in this shape language is considered to uniquely describe the shape of a word. The recognition of handwriting is subsequently implement by means of a logical comparison between the symbolic expression which represents an unknown word and the symbolic expressions which were built into the system during a manual training phase.

```
"million" ==> convex:concave:3(north:concave)
               :(north:LOOP):concave:(north:LOOP)
               :concave:north:concave:HOLE:2(convex:concave)
```

Later research in our own laboratory has shown that the human reader indeed pays special attention to the singular elements in the writing trace, such as crossings (Figure 4).

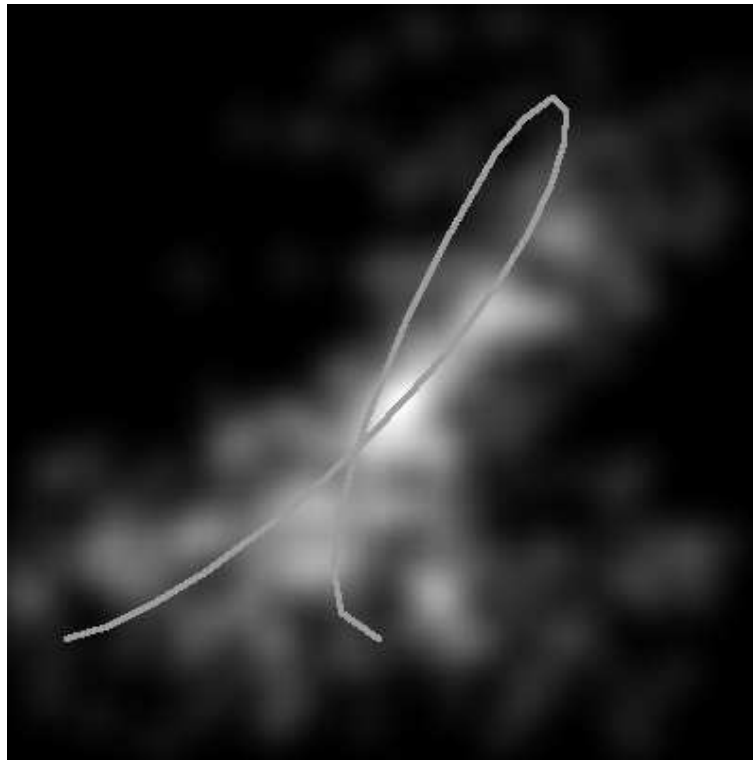


Figure 4: Density of selective attention in the human reader of cursive handwriting, superimposed with the shape of the trace (Schomaker & Segers, 1996). In this example, the crossing in the trace of the letter *l* draws the attention of a reader under time pressure.

The theory of Simon is elegant and the singularities can be detected easily with an algorithm. The word-search method itself is computationally costly but manageable in case of limited shape variation. In case of handwritings by a few writers and assuming a small lexicon the method worked reasonably well*, such that the idea emerged to apply this approach to automatic reading of handwritten amounts (courtesy amount) on the back side of French bank checks. The confrontation with reality was painful, however. Handwritings of only 20 subjects are completely insufficient to be able to generalize to the complete population of writers. A screen full of symbolic descriptions of the word *million* would cover only a tiny fraction of the possible shape variation in the population. The complexity of word-shape descriptions for all possible style is phenomenal and the computational load required for search in large symbolic data structures quickly become unmanageable, even for computers of today or of next year. The human pattern generator appears to be able to produce a sheer infinite variation of shapes in the two-dimensional plane.

*Word classification performance was 87% correct on the training set, 60% on the testset, using 25 word classes.

Figure 5 gives an impression of the variation in written forms (allographs) of individual letters. A random selection of letters “t” has been made from a large database.

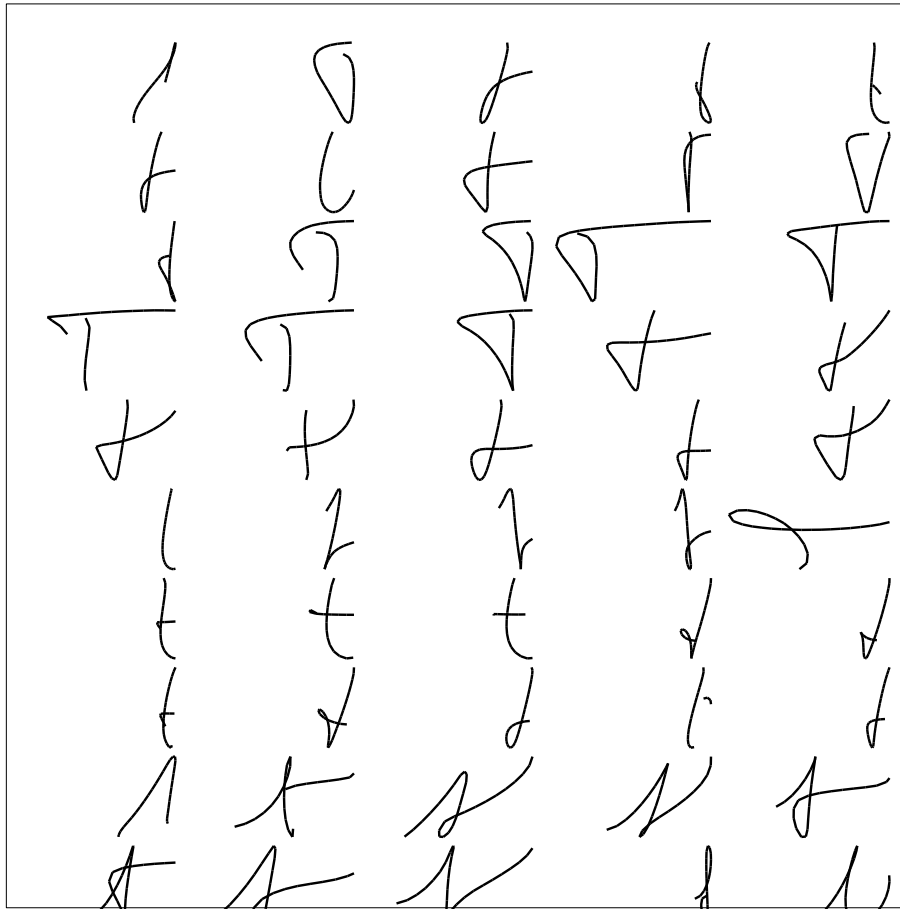


Figure 5: Allographs for the letter *t* (Vuurpijl & Schomaker, 1997)

The core problem of a logical-symbolic approach to pattern recognition resides in the fact that elements from sensory input are considered as “serious” at a very early stage of information processing. A symbolic identity is attached to small shape elements, while there still is uncertainty and massive variation. Once generated, a symbol takes part in formal operations and cannot be ignored, even if it concerns a spurious logical proposition. I therefore propose the following adage: “no noise more obnoxious than symbolic noise”. Whereas noise is a manageable concept in signal-processing theory, there exist no elegant noise-reduction or filtering methods in the symbolic world. Symbolic noise acts as pebble stones between the cogwheels of an engine. A wide range of experiences within the field of pattern recognition with the brittleness of a strictly symbolic paradigm stimulated a search for other methods.

This brings us to a second perspective on pattern recognition which stresses geometry and statistics rather than logic as the prominent tool for understanding and implementing perceptive systems. As was aptly posed by the Dutch physicist Koenderink (1990): “The Brain: (is) a geometry Engine”.

During the nineteen eighties and nineties of the last century, the field of pattern recognition enjoyed an accelerated development due to the evolution of existing ideas and the emergence of new methods: Markov modeling, artificial neural networks and Bayesian methods all yielded very interesting results in comparison with earlier approaches. We cannot handle these technologies in detail within the scope of this presentation. The essential common element of these methods is the fact that they are explicitly or implicitly of a statistical nature.

Instead of enforcing a symbolic formalism upon the natural data, the methods in this second group are based on the following assumption: “If there is regularity in the data, then there must exist an algorithm which

is able to detect this regularity". This adage was picked up by many researchers of my generation. I will give an example from the field of automatic script recognition.

According to a theory on the human writing process by the Handwriting Research Group of Nijmegen University since 1976, the movement of the pen tip can be segmented into strokes. Spectral analysis of pen-tip movement reveals a strong periodicity around 5 Hz. Measurement of the average duration of strokes subsequently shows that a mono-phasic movement in handwriting is delimited by minima in the absolute pen-tip velocity. A modal stroke in the writing process lasts about 100 ms. The vertical-velocity profile of two consecutive strokes in script approaches the shape of a whole-period sinusoid.

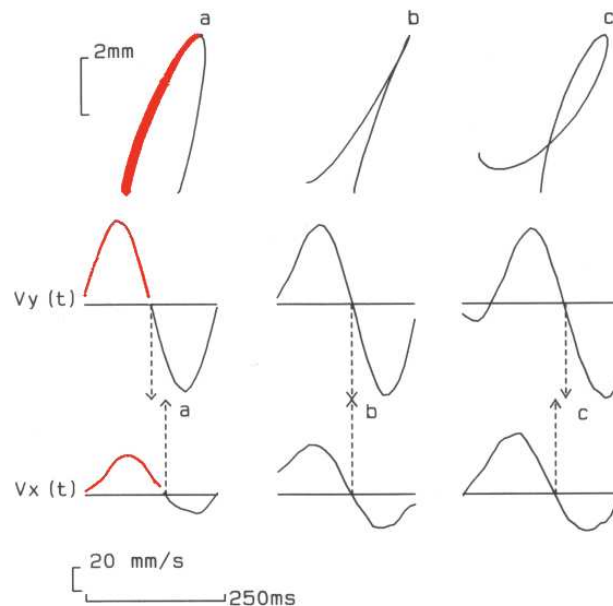


Figure 6: A stroke definition in handwriting based on pen-tip velocity and the phase difference between horizontal and vertical velocity. Three major shapes are shown here (strokes with bending points constitute a special case).

By specifying local phase differences in the movement, for the horizontal and vertical component separately, the large majority of velocity-based strokes in script can be modeled faithfully. The knowledge which has been gathered in this area is currently sufficiently detailed, such that it has become possible, within margins of uncertainty, to make an educated guess concerning local velocity and the order of strokes on the basis of a static image of the written product, e.g., the signature of Isaac Newton (Figure 7).

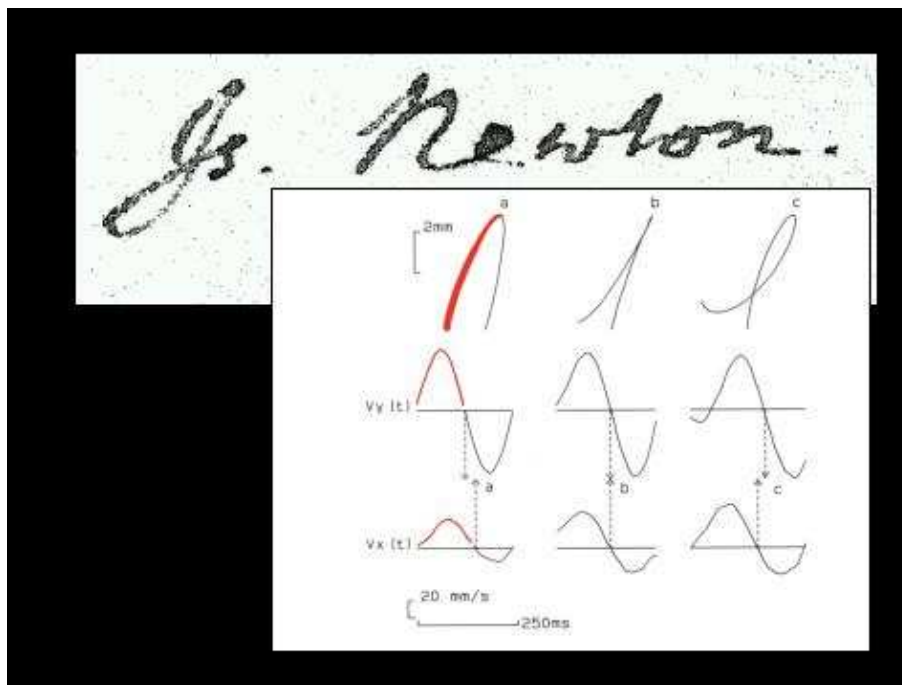


Figure 7: A local temporal reconstruction of Newton's signature is possible on the basis of existing theoretical knowledge, yielding reasonable estimates of speed and sequence.

On the basis of the accumulated body of knowledge on the writing process in the middle of the nineteen eighties, it was reasoned that such knowledge might be utilized in the recognition of writing movement, such as produced with a pen on an electronic tablet (XY digitizer).

After a number of experiments which were comparable to the structural-features method of prof. Simon, it was assumed that the upcoming neural network models could provide a solution to the problems of variation and variability of script. A powerful method which illustrates elegantly how regularity in data can be detected autonomously was developed in the late eighties by Teuvo Kohonen (1995) and was quickly introduced into the Nijmegen lab by my esteemed colleague Piero Morasso. This method is called the self-organizing feature map. Instead of enforcing a symbolic order on raw data, tens of thousands of isolated pen strokes are segmented from a large database and presented to such a self-organizing map containing only a limited amount of "cells". The goal is to obtain a map with prototypical strokes which describes the statistical structure of the lump of raw data in an optimal manner, given the constrained map size and dimensionality. Figure 8 shows an example of a Kohonen self-organized map of strokes.

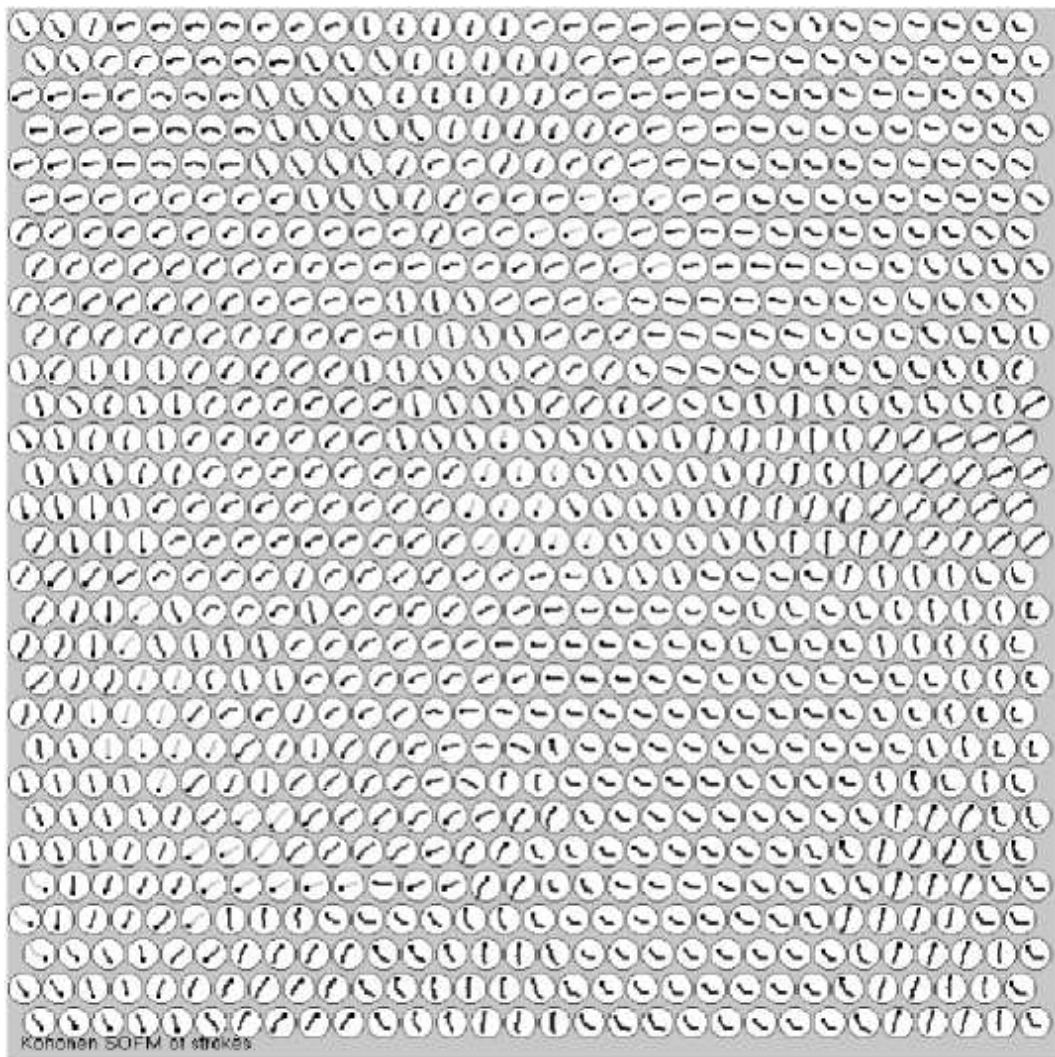


Figure 8: A Kohonen self-organized map containing a stroke-alphabet of script of a large number of writers. By using an excess of 30x30 cells, there is a gradual change of shapes over the map. However, for cursive-script recognition, a much smaller network of 400 or 625 cells would suffice (Schomaker, 1993).

On the basis of a stroke map, a written word can be represented as a path in this new, quantized space. The resulting transition network provides the basis for letter and word completion. The resulting system is functional and has been demonstrated at a number of occasions, such as a setup in the Scription museum in Tilburg, for several months. Although the performance of this system was sufficient to please the museum visitors, it remains a difficult fundamental problem to process the script of a writer who is completely new to the system. By using a number of complementary approaches in a multi-agent system, we have tried to incorporate a wide coverage of writing styles into this system.

The tool kit of pattern recognition is richly filled, currently. Apart from the aforementioned (hidden) Markov models and neural networks a new and powerful method has been added to the arsenal in the early nineteen nineties: The support-vector machine, which was developed by Guyon, Vapnik and other researchers of the former Bell Labs (1992). However, even this most powerful method for pattern classification does not offer a complete solution to the problem of reading machine in general. At this point it should be noted that in contemporary research, toy problems on the basis of only twenty writers have been left behind since long. In cooperation with a large number of companies and research labs we have collected a database of written samples for “on-line script recognition” in pen-based computers. This database contains more than 300k isolated words* and 450k isolated characters by more than 1000 writers. Although this amount is limited

*More than 23k different words

in comparison to similar public databases for speech recognition, the diversity which is present in this database poses serious problems for current research, which is progressing only slowly at the global scale.

Here, I would like to attempt to clarify at what points the problems are felt in the most painful way. A central problem constitutes the transition from a metric and geometric world to the world of symbols.

Starting from a high-dimensional representation with sensory data, a selection and projection takes place to a lower-dimensional feature space. This transform is the first powerful 'trick' of a pattern-recognition system within the geometric paradigm. It is interesting to note that the biological neural substrate is ideally suited to perform such transforms by means of networks of neurons.

As a simplified example we take the classification of an unknown letter as belong to one of three classes a , u and d . Let us define the following two features: F1, the angle of the line piece between the left and right vertical maximum of the letter, and F2, the length of this line piece (Figure 9).

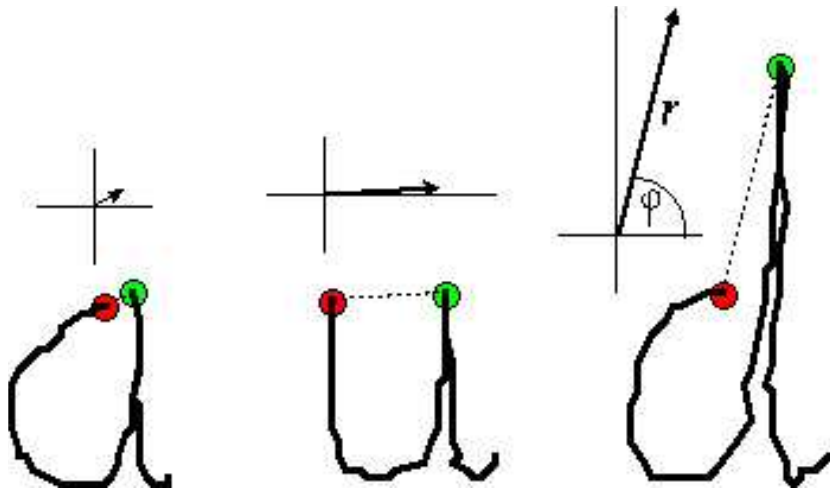


Figure 9: Useful features for the classification of the letters a , u and d : The angle and length of the line piece (vector) between the first and the second vertical maximum in the pen trajectory. These two features F1 and F2 allow for representing openness/closedness and the height of the letter d .

Based on about seven thousand examples of these letters the two-dimensional probability distribution of the features F1 and F2 can be determined. It appears to be a mountain landscape with three different peaks, one for each letter class (Figure 10).

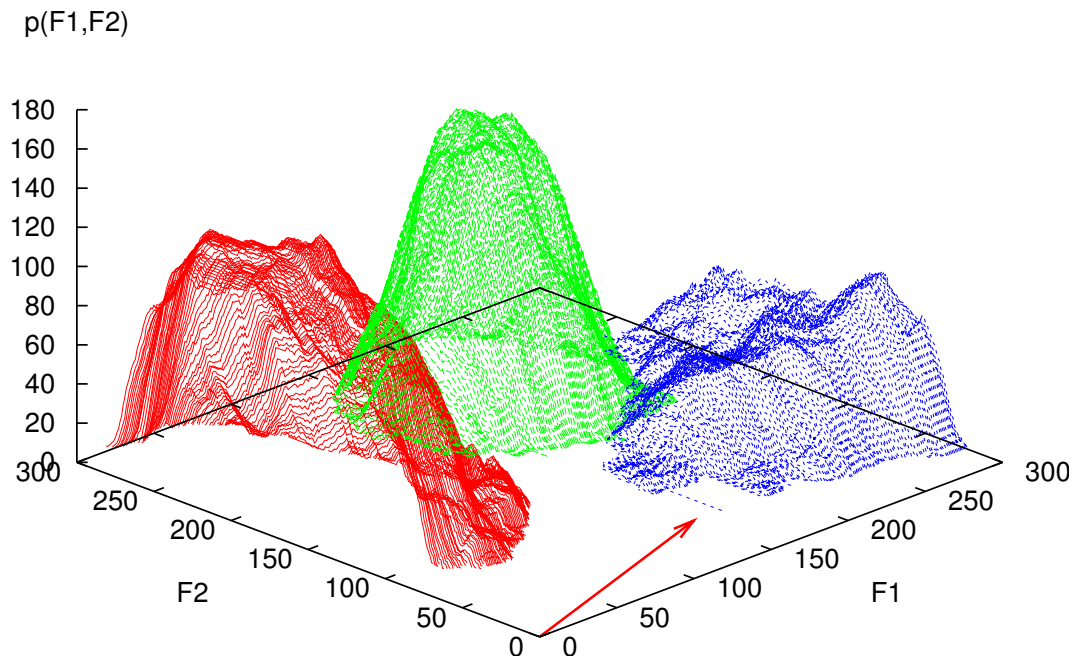


Figure 10: A 2-D probability distribution of features $F_1 = \sin(\phi)$ en $F_2 = r$, colored according to letter class. Also see Figure 9. The units along the axes are arbitrary.

Apart from probabilities, however, there is another issue to deal with, in patter recognition. Since the choice for a class has consequences for an autonomous system operating in the real world, the resulting mountain landscape needs to be molded on the basis of a cost evaluation. Indeed, misclassification has different consequences in different real-life contexts. A missed forged signature obviously has more serious consequences for a bank, than an incorrectly recognized word has to the user of a palm computer while editing. The modulation of the 'probability landscape' is the second powerful geometrical 'trick' which is possible within geometric/statistical pattern recognition.

However, we cannot escape fate: The ultimate goal of the whole process is to attribute the unknown vector (drawn at the base of Figure 10) to one of the three letter classes a , u or d , i.e., elevating the process to the level of symbols (and logic).

There are many methods which allow for finding separation boundaries between classes in high-dimensional space. For our 2-D example, a class separation can be visualized in color, as has been done in Figure 11. Also the identity of the classes is given in the next figure.

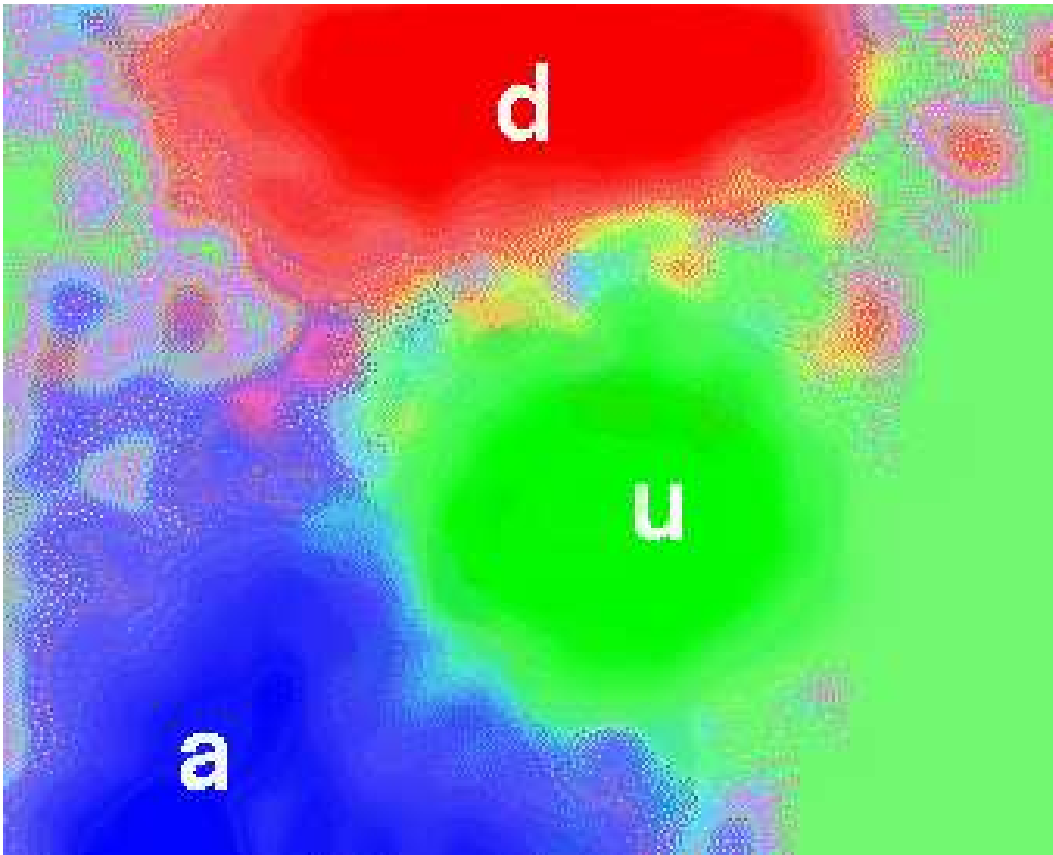


Figure 11: The attribution of letter classes *a*, *u* and *d* on the basis of the two features r (x-axis) and $\sin(\phi)$ (y-axis). In pattern recognition, there will always be problematic cases at the boundaries between the classes. Also, most methods make assumptions about feature-space regions in which no actual examples were ever observed.

In spite of the elegant possibilities of class-separation methods several problems remain. By choosing a class, there is always the possibility for misclassification. Also, it can be observed that the class boundaries may be of complex shape, locally, such that a perfect separation is not possible. Furthermore, most methods generalize (extrapolate), yielding system responses for feature combinations which were never observed in the training history. Such generalizations may work out well or cause havoc. The essence of these problems is that a choice, once made, cannot be modulated by the system on the basis of shape or cost evaluations: “no noise more obnoxious than symbolic noise”. An erroneous decision will be used in an unweighted manner and considered as ‘serious’ by means of symbolic input to all later reasoning stages. At the level of logic, “almost-a”s or “almost-u”s do not exist anymore.

Maybe the following actual example elucidates the problem more clearly. At an airport, two systems have been installed for the detection of weapons in luggage: A (electromagnetic) metal detector and a (sniffing) explosives detector. In case of a sufficient volume and mass of the sought metal, the metal detector will deliver an alarm signal. If there is a critical but subliminal perception of metal, the piece of luggage is allowed to continue, naturally. However, if this latter event is followed by an equally subliminal detection in the second system, i.e., the explosives detector, one would hope that an intelligent system will send an alarm, still, although both observations would be classified as harmless if occurring in isolation.

The consequence of the problem of deciding in uncertainty is that current systems need to postpone hard decisions as long as possible in the processing pipeline. At the same time, however, hard symbolic information constitutes the necessary basis for powerful reasoning mechanisms. In automatic script recognition it will be absolutely necessary to interpret character *shapes* within the framework of a context with expectancies concerning the *content* of the written text.

In this manner, both incommensurables, i.e., geometry and logic, are forcibly merged into an intertwined architecture which is made to function by a legion of system developers and programmers. The resulting hybrid contraption, however, is of a notoriously rigid nature. For each new application of automatic script recognition an immense amount of human mental effort must be spent. The technology which is used in an isolated digit classifier for bank check readers or postal address readers is virtually useless when it comes to reading notes from a pen computer, or reading text by camera in the 3-D environment such as a street. For each application, a specialized and inflexible system must be constructed. This rather unintelligent state of affairs does not only occur in automatic script recognition: Similar problems of comparable scope are encountered by those who are implementing systems for automatic speech recognition.

The autonomous perceptive and reasoning machine remains a dream.

[the Epilogue of this speech can be found in the Dutch text version]

References

George Boole The Calculus of Logic The Cambridge and Dublin Mathematical Journal Vol III, 1848 pp. 183-198

Guyon, I., Boser, B. & Vapnik, V.N. (1992). *A Training Algorithm for Optimal Margin Classifiers*, Proc. of the 5th annual workshop of computational learning theory, ACM, pp. 144–152.

Koenderink, J.J. (1990). The brain a geometry engine. *Psychological Research*, 52, 122-127.

Kohonen, T. (1995). *Self-Organizing Maps*, Springer Series in Information Sciences, Vol 30, Heidelberg: Springer.

Schomaker, L.R.B. (1993). Using Stroke- or Character-based Self-organizing Maps in the Recognition of On-line, Connected Cursive Script. *Pattern Recognition*, 26(3), 443-450.

Vuurpijl, L. & Schomaker, L. (1997). *Finding structure in diversity: A hierarchical clustering method for the categorization of allographs in handwriting*, Proceedings of the Fourth International Conference on Document Analysis and Recognition, Piscataway (NJ): IEEE Computer Society, p. 387-393. ISBN 981-02-3084-2