The Inference of Identity of Source: Theory and Practice

Christophe Champod, assistant professor, Institut de Police Scientifique et de Criminologie, University de Lausanne, BCH, CH-1015 Lausanne-Dorigny (Switzerland). Tel. +4121 692 46 29 Fax. +4121 692 46 05 e-mail: christophe.champod@ipsc.unil.ch

Introduction

What distinguishes criminalistics from other sciences is that it not only aims to identify objects or persons as members of known classes, but to individualise them [1]. Individualisation is an essential part of forensic science, the daily reward in fields such as fingerprint, earprint, shoeprint, speaker recognition or handwriting. With the extensive use of DNA — probability based — evidence and the evolving requirements for the admissibility of scientific evidence in the United States of America, older identification fields are subject to more rigorous scrutiny and are under the pressure of a growing demand of scientific data [2-4]. The conclusion of Jonakait well resumes the position of some lawyers: "In other words, if Daubert is taken seriously, then much of forensic science is in serious trouble" [6, p. 2217].

The aim of this essay is to delineate a general inferential scheme of the individualisation process. This theoretical framework will allow an analysis of practices within some personal identification fields (handwriting, fingerprints and DNA). Finally proposals towards a harmonised approach to identification inference process will be made. It will enable the adoption of coherent validation schemes, as required by the scientific approach.

In accordance with the theme of this colloquium, this essay will focus only on human identification fields, however, the extension to other identification fields, such as shoeprints, toolmarks or firearms, suffers no restriction.

The identification inference process

Among identification fields, the term identification generally denotes individualisation. In literature, this problem of identity of source is often treated by reference to "class" (even "subclass") and "individual" or "unique" characteristics. Comparisons between a recovered mark and a known print that lead to agreement only in class characteristics (without significant differences) will end up with "groupl identification" conclusions. Only when sufficient agreement of individual characteristics are observed in conjunction with class characteristics, then positive identification or individualisation conclusions can be drawn. The problem of exclusion may be considered straightforward as one significant observation against the proposed hypothesis of common origin is sufficient to reject it. However the distinction between "class" and "individual" characteristics are only conventional ways of describing selectivity. To avoid misidentification, the task of distinguishing among these categories
is essential, but the problem of inferring identity of source is more complex than this simple dichotomy. The key question relates to the concept of "sufficient agreement" between the questioned mark and the control sample allowing a conclusion of identity. Philosophically, identity of source cannot be known with certainty, and therefore must be inferred. As Kwan has demonstrated, a hypothetical-deductive method (assisted by methods of statistical inference) provides a reasonable explanation of how criminalists proceed in inferring identity of source [9]. The identification process can be understood as a process of reduction from an initial population to a restricted class or ultimately to a single individual. Two factors therefore enter into combination: (1) a relevant population of persons defined by size (and/or other particularities); in other words, each member of this population can be seen as a possible source; (2) a reduction factor based upon the combination of concordant characteristics of determined selectivity. In fact the reduction power is proportional to the rarity of the observed characteristics in the population under consideration. Of course, when individualisation is the goal, the object must be defined by a unique set of properties (a set that can be attributed to no other source).

The size of the relevant population can be conceived as an open framework or as a closed framework. By open framework, it is meant that the population at large comes under consideration, for example, all living persons on earth are initially considered potential sources. The closed framework corresponds to a situation where the initial number of persons is restricted to a specified set of suspected sources (for example by taking into account other evidence limiting the putative sources).

The reduction factor is probabilistic in essence, based on objective empirical data and/or on subjective evaluations related to the examiner's experience. DNA evidence for example relies on hard statistical data, whereas other identification fields such as fingerprints, earprints or handwriting rely mainly on educated assessments of the specificity of the shared features.

The identification process (either in open or closed framework) remains a narrowing-down process, reducing the number of possible sources or hypotheses. Showing that all alternative hypotheses that could explain the phenomenon at hand are excluded will validate the hypothesis that a designated suspect is the source. As the process is probabilistic, we need a tool to understand the reduction process, to show how a new piece of evidence will modify our prior degree of belief about identity. This logical tool is provided by the Bayesian approach that will be briefly introduce. The application of Bayes' theorem to identification evidence is not new. Advocated in 1904 by Henri Poincaré and his colleagues following the miscarriage of justice in the Dreyfus case (handwriting examination) [14], the approach has been applied by Finkelstein and Fairley [15] to other identification evidence such as bloodstains or fingerprints. This logical tool allows an appreciation of how events interact mathematically in a frame of hypotheses or propositions — hereinafter noted \( I \) and \( \mathcal{E} \). It clarifies the position (and duties) of the scientist as well as that of the judge and defines their relationship.

The Bayesian approach for identification evidence

The following terms have to be defined:
Background information that has been collected prior to the forensic examination. For example, data from a police investigation, eyewitness statements or data from the suspect's criminal record will typically contribute to I. This information will allow an estimation of the number of potential suspects or objects that could lie at the origin of marks.

E: The evidence: result of the comparison between a mark left on a scene and a person under examination.

ID: The mark has been produced by the person under examination.

? : The mark has not been produced by said person, and another unknown person is at the origin of the mark.

Proper definition of the two exclusive hypotheses (ID and ?) requires consideration of the context of the case: they are not always as straightforward or exhaustive as the above definitions might lead one to believe. As only the identity of source is concerned here, the hypotheses (or propositions) are framed at the source level as defined by Cook et al [16].

Bayes formula (1) below shows how prior odds on ID are modified by the evidence E to obtain posterior odds on the issue at hand (ID versus ?) using a simple multiplication by a likelihood ratio.

\[
\frac{\pi(x | ID)}{\pi(x | ?)} = \frac{\pi(x | ID) \pi(ID)}{\pi(x | ?) \pi(?)}
\]

(1)

Where:

\(\pi(x | ID)\) are the odds on identification prior to forensic examination according to the information (I) available.

The prior odds on ID are equal to the ratio \(\pi(x | ID) / \pi(x | ?)\).

The prior odds are thus related to the initial number of potential persons and the relative weight of the suspect regarding ID.

\(\pi(x | ID)\) are the odds on identification given the evidence E and the available information (I). The posterior odds on ID are equal to the ratio \(\pi(x | ID) / \pi(x | ?)\).

\(\pi(x | ID)\) is the probability of the shared features between the mark and the control material, given that the mark has been left by the person under examination (ID is true), and the relevant background information (I). This value is not systematically 1.

\(\pi(x | ?)\) equals the probability of the shared features
between the mark and the control material, given that
the mark has not been left by this person (\( \gamma \) is true),
and the relevant background information (\( \Omega \)).

The multiplication between the prior odds and the likelihood ratio to obtain
the posterior odds may be seen (provided the likelihood ratio is above
one) as the reduction process governing the identification process.

The estimation of prior odds will be based on the background information
(\( \Omega \)) gathered by the police on the case. Most of the time, these data are
unknown to the forensic examiner, hence the assessment of the size of the
relevant population should remain the province of the court (judges,
parties, members of the jury, etc.).

The scientific statement made to the Court by the forensic scientist
(whether statistical or subjective) is the exorision of the reduction factor
to be applied to the prior odds. In the absence of information on the prior
odds, it is not possible for the scientist to address the fact itself (the
probability that this particular person has produced this mark). He/she can
only address the degree of support given to this hypothesis (or
proposition) as opposed to the alternative. In the Bayesian approach, the
strength of the evidence is assessed by the probability of observing the
evidence under two chosen propositions. This ratio (called likelihood ratio)
has a numerator describing the adequacy between the mark and the
known print if ID is true and a denominator equal to the random match
occurrence of the shared features in the relevant population. The concept
of evidence is therefore relative: it reflects how observations should be
interpreted as evidence for as opposed to \( \overline{\Omega} \).

The posterior odds in favour of the identification itself — that the person in
examination has produced the marks — are a judgement based on prior
odds combined with the evidence. When the prior odds are considered as
outside the province of the scientist, the responsibility for the decision on
identification (posterior odds) remains the burden of the court.

The practice

Unfortunately, in practice, the identification process leading to
individualisation is generally operated in all fields within an open
framework, as defined for fingerprints for example: "[...] the determination
that two corresponding areas of friction skin impressions originated from
the same person to the exclusion of all others." [18, p.156]. Thus the size
of the population under consideration is systematically set for reasons that
remain obscure to its maximum. The adoption of this open framework
leads to conclusions on the issue itself either through positive
identification or with qualified (or corroborative) opinion (associations
qualified as possible, probable, etc.). When the identification evidence is
known to provide probabilistic values such as DNA, the practice remains,
as we will see, focused on random match probability of that particular DNA
profile.

Individualisation

The establishment of positive identification is an opinion: a statement of
probability expressing that the chance of observing on earth another person presenting the same characteristics is zero. No contrary evidence — even the most powerful alibi — will ever shake the expert’s certainty. According to Stoney, this highly subjective decision involves a ‘leap of faith’ [19]. Kingston noted that sometimes we become so accustomed to certain conclusions that we accept without questions that we tend to give these conclusions that status of certainty. However, positive identification remains an opinion, a ‘posit’ [20]. During the recent Daubert hearing on fingerprint evidence, the defense correctly qualified assertions of absolute certainty as inherently subjective and unscientific.

**Qualified (corroborative) opinion**

In some forensic fields such as fingerprints, practitioners have voluntarily excluded probability statements — other than exclusion and positive identification — from their conclusions. All pieces of evidence between these extremes are classified as ‘inconclusive’. The whole profession according to a resolution adopted by the IAI (International Association for Identification) in 1980 largely followed this point of view, known as the ‘positivity’. This policy is opposed to the scientific principles governing the interpretation of scientific evidence (23, 24). The evidential value cannot be so drastic, so clear-cut; an increasing scale runs from exclusion to identification. In fact, there is no logical reason to suppress grey levels between white (exclusion) and black (identification). Evidence becomes relevant when it tends to make the matter at issue more or less probable than otherwise (25), and even if there is no certainty, it should not be ignored. This leads us to the analysis of qualified opinions.

In many cases, a forensic scientist will not be able to provide a definitive answer but only a probabilistic opinion. If the ultimate set of specific features is not present or not detected in the evidence, then the examiner will not provide an identification but will express a probability statement, verbally or numerically, which attempts to assess the value of the evidence and reflects the statistical uncertainty. Such conclusions are common in handwriting examination (verbally) or DNA analysis (numerically). For a verbal scale, an agreement on the following terms seems to be achieved: identification, very probable, probable, possible, inconclusive (could not be determined), appears not, negative, not suitable for analysis.

What is the meaning of conclusions such as: ‘it is possible (or probable or highly likely) that this mark has been left by this particular person’? Such verbal statements are clearly expressions of posterior probability obtained in an open framework. It has been demonstrated that these terms are quite uniformly understood (in a numerical conversion) by experts or jurors (27):

<table>
<thead>
<tr>
<th>Verbal statement</th>
<th>Posterior probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive terms</td>
<td>0.999</td>
</tr>
<tr>
<td>likelihood bordering to certainty</td>
<td>0.98</td>
</tr>
<tr>
<td>Likely</td>
<td>0.85</td>
</tr>
<tr>
<td>----------------------</td>
<td>------</td>
</tr>
<tr>
<td>Very well possible</td>
<td>0.75</td>
</tr>
<tr>
<td>(plausible)</td>
<td></td>
</tr>
<tr>
<td>Possible (events)</td>
<td>0.60</td>
</tr>
<tr>
<td>Negative terms</td>
<td></td>
</tr>
<tr>
<td>Possible (events) not</td>
<td>0.40</td>
</tr>
<tr>
<td>Very well possible</td>
<td>0.25</td>
</tr>
<tr>
<td>(plausible) not</td>
<td></td>
</tr>
<tr>
<td>Not likely / unlikely</td>
<td>0.15</td>
</tr>
<tr>
<td>Highly (very) unlikely</td>
<td>0.02</td>
</tr>
<tr>
<td>Likelihood bordering certainty not</td>
<td>0.001</td>
</tr>
</tbody>
</table>

**Random match probability**

The assessment of DNA evidence is different as examiners rely on statistical data to provide the court with a random match probability of the DNA profile in a given population. However, in 1997 the FBI decided to suppress probabilistic statements in favour of a definitive answer like for fingerprint evidence when the statistical figures involved are small enough to suppress any reasonable degree of scientific uncertainty.

The above decision schemes and practices (individualisation and qualified opinion) lead to the following discussion when confronted with the Bayesian approach.

**Discussion of the practice**

**The prior odds**

When working in an open framework, the scientist (not the court!) sets the size of the relevant population to the total number of persons on earth. It seems illegitimate to set a priori the size of the relevant population at its maximum. Indeed, the number of potential sources, which could be at the origin of the mark, may be restricted by other evidence available (witness testimonies, other forensic evidence, etc.), each case bearing its own specificity left to the appreciation of the court. Presenting the evidence in an open framework is far too conservative, adopting systematically the extreme Defence attorney's position: trying to make the Court believe that all persons or objects on earth can be at the origin of the traces. Assessment of the prior odds should be case specific and as such may be viewed as outside the duties of a scientific expert, and should remain the province of the factfinder [13]. Consequently the open framework adopted in most identification fields forces the scientist (often without notice) to assess prior odds in a way that could be incompatible with the case under examination. Time has come to call into question this dogmatic approach.
The posterior odds

Any statement on posterior odds needs to consider both prior odds and a likelihood ratio. A conclusion of individualisation is first and foremost a statement on the posterior odds. In addition the decision implies something more: the application by the expert of a personal threshold (rounding the posterior probability on 0 to 1) which is in essence a qualification of the acceptable level of reasonable doubt. It is tantamount to a judgement on moral certainty. A conclusion of individualisation is then based on statistical probabilities and on an appreciation of the concept of "beyond a reasonable doubt" [29]. The danger of an error is considered so small that the philosophical limit that human beings usually observe on knowledge and truth simply evaporate. Recently, the FBI has adopted a policy allowing identification of biological traces to "a reasonable degree of scientific certainty" [30]. The procedure invites the scientist to assess the prior odds, to combine it with a 'frequency' and to fix the level of reasonable doubt. In addition to the questionable legitimacy of the assessment of prior odds already mentioned, would jurists accept that the concept of reasonable doubt on the identification of a suspect escapes their province and is imposed onto the court by the scientists? The response in the doctrine is negative, as expressed by statisticians [31, p. 141], legal scholars [12] or forensic scientists [32, 33]. Hence, the scientist should avoid usurping the role of the court by making assumptions about the prior odds and statements about the posterior odds on an issue. When deleting these components from the Bayes' theorem, only the likelihood ratio is left to the consideration of the forensic scientist.

Working blindly in an open framework is in favour of the suspect but does not necessarily serve justice. Indeed it leads to the following paradox when concerning corroborative identification evidence: applying Bayes' formula (1) in a open framework of 5 billion individuals, the likelihood ratio of $5 \times 10^{10}$ leads to the most negative corroborative verbal statement (likelihood bordering certainty not with a posterior probability of 0.9991). Such evidence however represents very strong support for the identification. For the court, all pieces of evidence that lead to such reduction factors will be considered as highly relevant (by analogy with DNA evidence for example). However, the verbal statement does not make this evidential value very clear. To escape this paradox and to avoid assessing prior probabilities (and corollary posterior probabilities), the scientists could simply remain focused on the likelihood ratio.

In the light of what precedes, it would be best if most identification fields re-examined their practice of assessing prior and concluding on posterior odds. The future lies in an assessment of forensic evidence in the perspective of its likelihood ratio, as it constitutes a coherent way of describing objectively the weight of scientific evidence. This recommendation may be seen as a complete questioning of the role of the expert witness; aren't the questions always submitted in the form: has this mark been (possibly, probably, with certainty, etc.) left by John Doe? This question is the one the court wishes the scientist to answer, but we just have explored how such conclusions carry more than just scientific examination, discovered the fine line that delineates science from opinion. I am not advocating the banishment of all statements on posterior odds from expert reports. It is up to every examiner to decide if the scientist decides to get involved in the assessment of the credit or be attributed to
one hypothesis or the other (statement on the issue itself), he/she — as well as the factfinder — must be fully aware of the additional step taken. The statement report should imply or cautiously avoid the premise possibility (about prior probabilities) which, combined with the evidence, permits the expert to draw this conclusion. Here the expert is more than a scientist, he/she is a witness giving a personal opinion based on specific knowledge (scientific and non-scientific) of the case. The report (or testimony) must be univocal in this respect.

The likelihood ratio

The likelihood ratio necessary for an individualisation to be declared is either infinite or so high that it leads to posterior odds that prevent any discussion about doubts. From a numerical point of view, if we accept that the posterior probabilities must be below a certain threshold value in order to declare an identification (i.e. a modest value of 0.9998), then the likelihood ratio that must be considered to achieve such a pre-set value in an open framework of 5 billion persons must be equal to $2.5 \times 10^{12}$ which represents about 5000 times the size of the initial population of 5 billion. A conclusion of individualisation (according to this threshold and in an open framework) is the expression of such a high reduction factor. The question is to know if an expert can derive from the examination of identification evidence such likelihood ratios or if such high likelihood ratios can even be conceivable.

We will concentrate our attention on the fingerprint field as it represents one of the more recognised ‘scientific’ evidence leading to definitive demonstration of identity. It is axiomatic that any two items human beings may be differentiated, provided the use of sufficiently collected analysis. But is the latent fingerprint examiner in possession of such an ultimate selectivity when the material under examination (recovered or known) is limited, partial or degraded? The flux of the matter is not really the individuality of the friction ridges but the ability to derive such a persuasive likelihood ratio from a small-distorted latent fingerprint fragment, revealing only a small number of basic ridge characteristics.

We have underlined the subjective nature of fingerprint comparison and evaluation. Kingston suggested two ways of assessing such subjective educated appraisals: one is through proficiency testing, the other is to assess how well subjective estimates correspond to objective ones [20]. Practice has shown that erroneous identification may be declared either in real practice or during collaborative testing. Moreover, when fingerprint specialists are asked to examine the same sets of comparison, large differences can be recorded either in the number of concordant minutiae [34] or in their ultimate conclusions. Hence, when examiners are tested as black boxes, the outcome does not match their peremptory attitude. Regarding isolated subjective estimates, Osterburg and Bloomingston showed how the assessment of the significance of a given type of minutiae varies among examiners [35]. Whereas, the fingerprint field is often presented as a science (the FBI tutorial is indeed entitled “The Science of Fingerprints”), as soon as we inquire into the probabilistic model that allow deriving adequate probability estimates, we fail to find coherent answers. As Stoney pointed out, none of the proposed models [36], even the most recent ones [36, 37], have been subjected to empirical validation. Indeed, inasmuch as a statistical method would suggest qualified (non-absolute) opinions, the models are on principle rejected by
the fingerprint profession invoking the 1980 Al resolution [38, p.72].

The recognition of identification fields as scientific domains seems deeply related to the capability of the field to provide reliable statistical estimates (either objective or subjective) of the rarity of identification features. This need was recognised by eminent criminalists i.e. for fingerprints [39] and for handwriting [1] but, more than thirty five years after, very few studies have gone along these lines. Interestingly, the concrete request for scientific inquiry into the nature of the identification process has come from the courts instead of the experts.

Recently, following Daubert and in the light of critical papers published in law journals [41, 42], the scientific status of handwriting comparison has been reviewed in various cases, leading to its admission as a non-scientific skilled testimony (the most famous case is United States v. Starzegczyl ). These decisions have promoted a scientific reanalysis of the field (various peer-reviewed papers attest this trend) and the publication of extensive treatises on this matter [43].

It is mostly the pressure from the court, more than the quest for science, that promotes research in forensics identification fields. It is up to the scientists to reverse this trend and initiate scientific studies dedicated to the identification process especially in fields such as fingerprints, earprints or speaker recognition.

The way forward

The Bayesian approach provides identification evidence with a logical structure, allowing the scientist to understand the crucial difference between assessing the weight of the evidence and assessing the issue in front of the court. As far as the weight of the evidence is solely concerned, the likelihood ratio-based approach is the best available model as it gives the principles that should govern the interpretation of evidence [26]: (1) interpretation takes place within a framework of circumstances (f); (2) interpretation needs the consideration of at least two propositions (here A and B) and (3) it is necessary to consider the probability of the evidence (E) given each of the stated propositions.

These principles are encapsulated into the likelihood ratio (formula 2) which forces the scientist to consider the value of the evidence under two competitive hypotheses provided by the court and not on the value of the issues themselves.

\[ \frac{P(E|A)}{P(E|\neg A)} \]  

(2)

The denominator of the likelihood ratio is best estimated through a probability of a random match of the concordant features in the relevant population. The relevant population is dictated by the alternative hypothesis considered and by background information related to the case. The numerator of the likelihood ratio asks for the probability of the evidence if the suspect has left the recovered evidence. This probability is not systematically equal to one and must be assessed in each case taking into account the intra-variability of the whole process that generates the mark. I have always been surprised by the facility with which examiners
find reasonable explanations for discrepancies and assume (without clear elicitation) that the numerator of the likelihood ratio to be simply one. This tendency is observed among fingerprints examiners [44] but also among DNA scientists [45] and other forensic scientists. However, we cannot assess the value of the evidence without carefully addressing the numerator of the likelihood ratio. Hence, it is recommended not to only focus on the denominator (the random match probability), as suggested by the dominant DNA reporting practice. Moreover, the random match probability, when presented alone, suffers many drawbacks such as its risk to be misinterpreted by scientists or members of the court (i.e. the well-known prosecutor’s fallacy) [46] and its incapacity to deal with complex evidence such as mixed DNA samples or cases involving multiple traces or offenders. For these reasons, it is strongly recommended to present the evidence in a likelihood ratio format [47].

Statistics definitely play a role in the assessment of evidence and its future for identification fields can take the following paths:

1. Use statistical data obtained through survey among relevant identification features to calibrate experts' subjective judgements.

2. Calculate likelihood ratios through adequate modelling either using existing technologies or following new developments. Existing forensic identification databases (such as AFIS systems or SCRIPT) are largely under exploited in generating evidence. The efficiency of the match algorithm of existing computer-based searching systems could be used to assess the likelihood ratio associated with a match value (MV). To obtain such figures, one would need to calculate the ratio between the probability densities for each match value when considered under proposition α or under proposition β. Because of the well-known efficiency of such systems, the likelihood ratios can be expected to be quite high.

These lines of research will make extensive use of a model of the reality, which tends to assess the intra- and inter-variability of marks. This model will generally represent only part of the available information (typically for fingerprints, quantitative data such as the number, type and orientation of minutiae) used by examiners to evaluate a comparison; it will (at present at least) hardly incorporate qualitative factors (such as ridgeology). The comparison process (pattern recognition which results from complex interactions between eyes and brain) is more complicated than any statistical model could propose today, even assisted by a computer. But systematic approach and computer offer the ability to search among huge databases and systematically compare pairs of non-matching items conferring to this approach superior scientific credentials than any examiner's casework experience. Again, the likelihood ratio derived using this process has to be seen as an element entering into the decision scheme of an expert. Both approaches (based on statistical data and subjective) are complementary and could even lead to the design of knowledge based systems helping the examiner in making decision through adequate procedures, questionings and data disclosure coming from statistical studies or the knowledge of past cases. The identification fields have been urged to present a more rigorous formalisation of the individualisation process, I share the opinion that a probabilistic approach
is part of the answer.

Acknowledgments

The author wishes to thank Tacha Hicks and Cédric Neumann for commenting on the manuscript.

Bibliography


