

# Computer Tools for Genealogy Reconstruction\*

Carlos Diuk<sup>†</sup> Enrique Tandeter<sup>‡</sup>

August 16, 2001

## Abstract

This paper deals with a set of computer tools developed for reconstructing a genealogy from a collection of marriage records. The tools were successfully applied in constructing genealogies of inhabitants of the Indian villages of Sacaca and Acasio in present-day Bolivia, for the 120-year period extending from 1690 to 1810. As source we used the 11750 marriage records from the period.

The process comprised six stages, all of them supported by an automatic computer-based method or by a manual method, aided by computer tools for handling and visualizing the data.

The six stages involved were:

- Orthographical homogenization and grouping of names. Through the use of a positional sub string equivalence computer dictionary, different spellings of equivalent names were grouped together. Other names were further grouped or separated manually, with the assistance of Andean linguistics specialists.
- Coding of other associated data, like *ayllus* (sub-groupings) and residence information.
- Record linkage was then used to identify repeated appearances of individuals in the database. Possible types or instances of re-appearance were selected, and different and complex record linkage criteria were defined for each case. Some complex full-name equivalence criteria were also defined, especially in the case of women's names.
- Automatic consistency checks were defined and run. Most of the inconsistencies were manually resolved, with the help of computer tools specifically designed.
- The genealogy was generated in the format required by GENOS, a software package developed at the Collège de France for the analysis of kinship.
- More complex consistency checks were run against the final genealogy, and inconsistencies were again manually solved.

---

\*This work is part of a project in which Luis Acosta, Marisa M. Diaz, Raquel Gil Montero, Maria E. Immolesi and Ana M. Presta take part. We would like to give our very special thanks to Dr. Pablo Jacovkis (Facultad de Ciencias Exactas y Naturales - UBA) for his comments and suggestions for this paper. We would also like to thank R.S.Davies, Cambridge Group for the History of Population and Social Structure, and Laurent S. Barry, Laboratoire d'Anthropologie Sociale, Collège de France, for their comments along the process.

<sup>†</sup>Dep. de Ciencias de la Computación, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Argentina

<sup>‡</sup>CONICET/PROHAL (Programa de Historia de América Latina), Instituto de Historia Argentina y Americana "Dr. Emilio Ravignani", Facultad de Filosofía y Letras, Universidad de Buenos Aires, Argentina

# 1 Introduction

There have been many successful record linkage and family reconstitution projects in the past, but mainly involving European populations. Most of them were based on parish records (baptisms, death and marriage records), population and economic census, or election polls.

We have found no successful attempts to do massive record linkage based on sources from colonial Latin American populations, no matter the period of time.

Our project attempts to reconstruct as fully as possible Indian marriage strategies in terms of kinship, based on statistically significant sources. Thus, we try to establish genealogies, the longest possible in time, that would reveal the largest quantity of existing links, both of consanguinity and affinity, between marriage partners.

We begin with a digitalized source consisting of the 11750 marriage records from the indigenous villages of Sacaca and Acasio, in present-day Bolivia.

We assume that the source is a complete one, registering every marriage which took place during the 120-year period extending from 1690 to 1810. This assumption is sustained by a nuptiality estimation we made based on the population estimate for the villages by Boleda - Tandeter [8].

The first obstacle, inherent to this kind of record linkage projects, has to do with the standardization and homogenization of names.

There are different kinds of difficulties to cope with:

- Orthographical variations, due to changes in the use of names or simple variations made by the recording parish priest.
- Typing or interpretation mistakes during data entry. The source is a micro-filmed version of manuscript records, that was manually digitalized.

Known standardization methods were examined and tested, with poor results. For example, *Soundex* was used as a first approach, but revealed itself as not suitable for Spanish or Indian names, being an algorithm thought in and for English names. Morris [5] mentions "the need to develop other systems according to the cultural and language base of the records concerned".

Our approach was a tailor-made rule-based process, with a final manual revision and re-grouping of names into standards with the assistance of Andean linguistics specialists. **Section 2** describes this process.

Next, other associated data was also standardized, so that the process of record linkage could now be done over a completely standardized corpus. In **Section 3** we describe this standardization.

In **Sections 3** and **4** we give a detailed description of the contents of each record, but essentially they include the names of the bride and groom and, in the case of bachelors, the names of their parents. The constructed genealogy will be thus extended in two ways:

- By identifying a reappearance of a bride or groom as parent in another record, we add 1 degree of depth to the line.
- By identifying a reappearance of parents, we find brothers and sisters, thus extending the genealogy laterally.

- Identification of a bride or groom in a second marriage lets us identify re-marriage occurrences.

**Section 4** describes this process of record linkage in detail.

In **Section 5**, we describe how we dealt with the inconsistent results that record linkage yielded.

Once we finish the identification of the database, we will process the resulting genealogy at the Social Anthropology Lab of the Collège de France in Paris, where a specific computer software called GENOS has been developed by Laurent S. Barry in order to analyze the marriage strategies in terms of kinship relationships of consanguinity and affinity between the marrying parties.

We finally believe that this organization and break-up of the record linkage process into six clear stages, and the tools developed for each stage, are useful and general enough so as to cope with very different record linkage requirements.

## 2 Name standardization (Stage 1)

We use a rule-based process for name homogenization, where rules are defined as a positional sub-string equivalence dictionary.

This computer-based dictionary is made of groups of strings that we define as equivalent, noting not only the string itself but also the position in which it should appear in a term to be considered part of the group.

To begin with, we organize all the names and surnames which appear anywhere on our database as a dictionary of independent terms, obtaining in our case a list of more than 6000.

The historian is now presented with a software tool which lets him/her define and test rules.

Figure 1 shows the user interface for the dictionary definition software tool, where you can see an example of an equivalence rule.

The interface shows the following rule definition:

**Rule 1** *GUA* at the beginning of a term equals *HUA*

The historian creates, modifies and deletes rules of this kind, and an option is available to test each new rule against the corpus of terms (the 6000 names and surnames in our case). The test simply shows all the terms that would be affected by the rule being edited, and the new spelling variations that would derive from its application.

The dictionary for the Sacaca and Acasio database is made up of 141 rules of this kind.

Consider now the following -additional- rules from our system:

**Rule 2** *H* at the beginning of a term equals  $\lambda$  (*nothing*)

**Rule 3** *PA* at the end of a term equals *BA* at the end of a term

We will show how an actual set of terms was grouped together based on the rules above.

The process begins by seeking, for each of the sub-strings in our dictionary of rules, its appearance in any of the terms in the term dictionary. If it finds it -in its

correct position-, a new term is added to the corpus, which is made up of the original term with the string being sought replaced by every other string of its equivalence group. The process keeps track of the original term from which the new terms are created.

This way, the table of terms keeps growing, and each rule is not only applied to the original terms but also to the ones generated by the process. But as this process might never stop, with the set of spelling variations growing infinitely, we made an arbitrary decision of repeating it 5 times. We base this decision on the expectance that two equivalent terms will not have more than 5 spelling variations among them.

Here's how the first of the 5 passes would work starting with the term *HUAMPA* (a real term in our database). Just suppose that our rule dictionary is only composed of the three rules mentioned above (instead of the actual 141):

The original term *HUAMPA* is affected by Rule 1 (starting  $GUA \equiv HUA$ ), so the corpus of terms now becomes: *HUAMPA GUAMPA*  
 ... the corpus is now affected by Rule 2 (starting  $H \equiv \lambda$ ), becoming: *HUAMPA GUAMPA UAMPA*  
 ... rule 3 (final  $PA \equiv BA$ ) gives us: *HUAMPA GUAMPA UAMPA HUAMBA GUAMBA UAMBA*  
 ... and the process goes on with pass 2 of 5, successively applying every rule to every term in the corpus

The final step is a manual revision of the grouped -and non-grouped- terms. A new software tool allows the historian to re-group, separate and create new groups. Andean linguistics specialists were consulted here. Figure 2 shows the user interface for this tool.

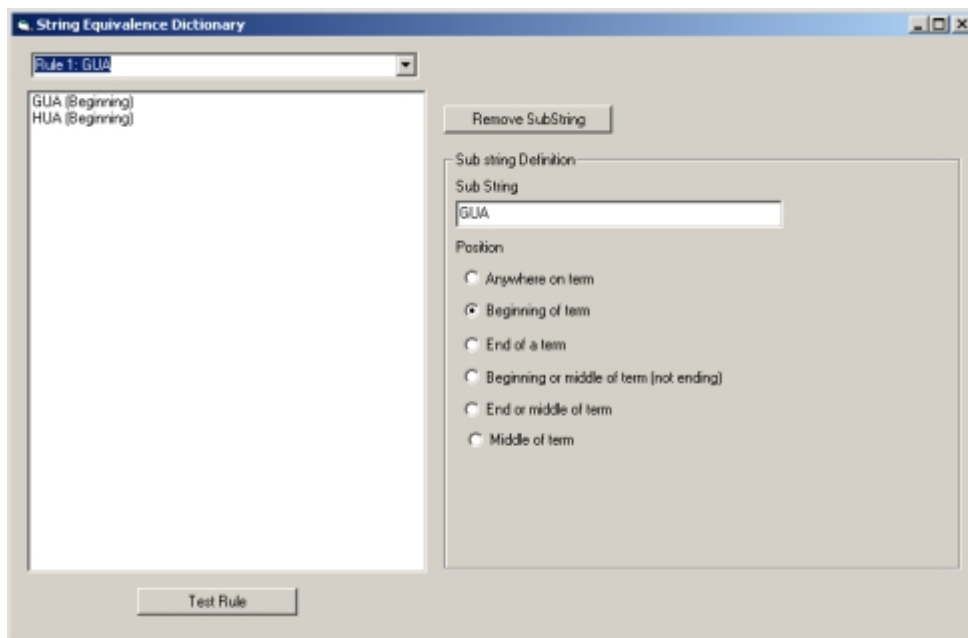


Figure 1: Equivalence dictionary rule-definition user interface.

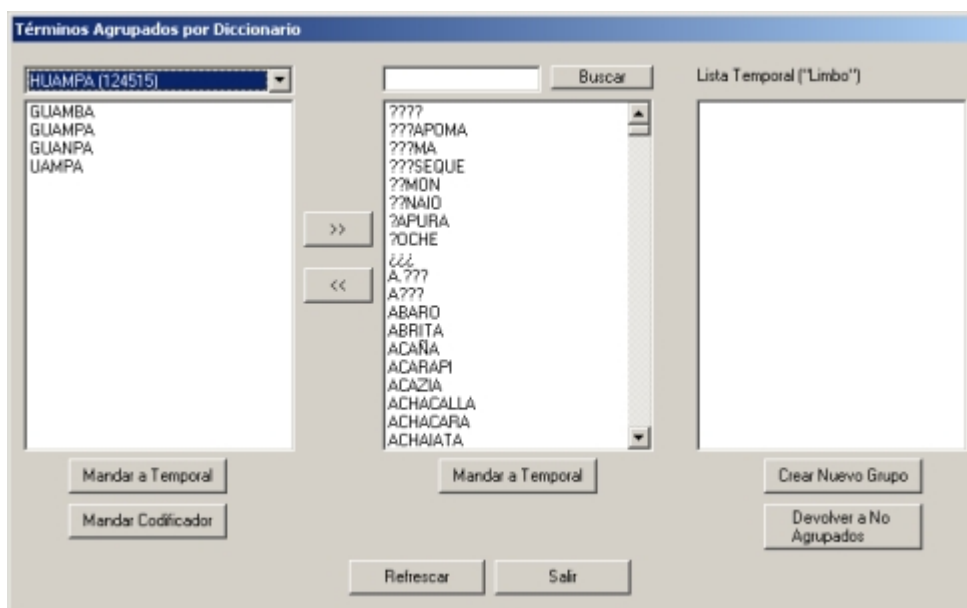


Figure 2: Software tool for re-grouping and separation of groups of terms. The left column lets you see the contents of each group. In the middle column, non-grouped terms are listed, and they can be added to the group on the left column. The right column is used for creating new groups, based on one or more groups, plus non-grouped terms.

When the process finishes, a code term is selected from each group (suppose it's *UAMPA* in our example).

Next, we generate a copy of our source database, replacing every appearance of a term with the code term from its group.

From now on we'll be working with a codified database. Following our example, people whose surnames were *HUAMPA*, *GUAMBA*, *GUAMPA* and *UAMPA* in our database are all now coded as *UAMPA*, and will thus be considered as possible identifications in the record linkage process.

In the Sacaca and Acasio database, from the 6000 original terms, 1400 remained uncoded, while the other 4600 were grouped into 896 groups of equivalent terms, resulting in an average of 5.13 terms per group.

Another approach to this problem, incorporating a probability measure into a rule-based method, can be found in Bloothoof [2].

### 3 Codification of other associated data (Stage 2)

Each marriage record contains, besides the names of the people involved (as described in the introduction), the following associated data:

- Bride and groom's marital status.
- Bride and groom's *ayllu*<sup>1</sup>.

<sup>1</sup>Andean kin group that generally traced its descent to a common real or mythical ancestor and

- Bride and groom’s residence information.
- Bride and groom’s origin.
- Bride and groom’s fiscal or ethnic category.
- Parish where the marriage took place.

In order to run a record linkage process over the database, involving criteria which takes into account some of this information, it must first be standardized.

As opposed to what happened with names, there were few variations for each field, although in some cases they had to be standardized taking into account not only orthographical variations but also historical knowledge or simple common sense. This part of the process was completed by hand.

If it were the case that we had too many variations and different information for a field, a rule-based process similar to the one used for names could have been used here. In the case of *ayllus*, for example, even though they were standardized manually, the same rules used for names would have worked.

Besides these usual standardization problems, a new requirement was found here: *ayllus* are grouped into a larger sub-grouping called a *moiety*. Sometimes, not the *ayllu* but its *moiety* are recorded in the source, and we didn’t want to lose this information. The same person might appear once with his/her *ayllu* stated, and somewhere else just with his/her corresponding *moiety*. Although the matching would be weaker, it is still a possibly valid linkage.

To deal with this, *ayllus* were related to their corresponding *moieties*, and used as criteria for record linkage.

In the Sacaca and Acasio database, there were 229 spelling variation for *ayllus*, which were coded into 123 actually different ones. 10 of those actually correspond to the villages of Sacaca and Acasio, and we thus know which of the two *moieties* of these villages correspond to each *ayllu*. For the remaining 113, corresponding to other villages from which the brides and grooms had migrated, we don’t have *moiety* information. Anyway, only less than 2% of brides and grooms who have *ayllu* information belong to one of these 113.

## 4 Automatic Record Linkage (Stage 3)

The main stage of the process is now the automatic record linkage stage. We seek to link individuals with possible re-appearances of themselves in other records of the same database.

Each of our records includes the following information:

- Groom’s first name and surname.
- Bride’s first name and surname.
- Groom’s parents’ first names and surnames.
- Bride’s parents’ first names and surnames.

---

had collective rights to land.

- Groom’s former wife’s first name and surname (in case of widowers).
- Bride’s former husband’s first name and surname (for widows).
- Groom’s *ayllu*.
- Bride’s *ayllu*.
- Other residence, origin and related information, as detailed in section 3.

Individuals getting married may reappear when they remarry, when their spouse does (in the "former bride/groom" field), or when their children get married (as parents).

So we first identify all linkable pairs of fields, and we then establish the linkage criteria for each case. We decided to define linkage criteria based on couples instead of individuals, as no individual would be possibly identified in different records without identification of his/her spouse. This way, 12 possible linkage cases were selected, shown below :

<i>Couple</i>	<i>Groom's parents</i> <i>Bride's parents</i> <i>Groom – former bride</i> <i>Bride – former groom</i>
<i>Groom's parents</i>	<i>Groom's parents</i> <i>Bride's parents</i> <i>Groom – former bride</i> <i>Bride – former groom</i>
<i>Bride's parents</i>	<i>Bride's parents</i> <i>Groom's parents</i> <i>Groom – former bride</i> <i>Bride – former groom</i>

The process of defining adequate criteria for deciding whether two records should be linked through any of these linkage cases always involves a compromise between a tempting laxity and a sometimes-disappointing strictness. Too lax criteria will spawn lots of possible linkages, but necessarily involving many false identifications, which will lead to numerous inconsistencies and ill-supported hypotheses. Too strict criteria leave out many identifications.

On a first approach to the problem, what we actually did was assign a positive value to each field coincidence among two records, and in some cases negative values for differing elements. This let us define a *weight* function, a quantification of the probability of two records containing references to the same person<sup>2</sup>.

As an example, a first or surname coincidence would sum a value of 1 to the *weight*, *ayllu* would weigh 0.5, and other data such as origin or residence 0.25, or even 0.125. Also, an *ayllu* contradiction would have a -0.5 value, and lack of *ayllu* information would weigh 0. Some contradictions in not too relevant data, such as residence information, would not have a negative impact.

The result of this method was a set of pairs of records, with the *weight* value for their probability of containing an identification. For these pairs, we then chose an

<sup>2</sup>Properly normalized to a  $[0 \dots 1]$  co-domain, the *weight* function can be regarded as an actual probability function.

*acceptance* threshold, a *rejection* one, and what we call a *gray region*, with pairs of records to be checked manually.

However, this process proved to be a faulty one. It's not easy to choose the positive and negative weights for each field match, and suitable thresholds so that results are trustworthy enough and the *gray region* is not absurdly large. A sum of small positive weights may cover for an important negative value, which is sometimes key in the identification. And making these key weights larger, or reducing the weight for side information, drops many identifications into the *gray region*, or accepts faulty ones.

Some general conclusions were drawn out from this first attempt:

- Residence, origin, and other related information is too weak, and often not registered, so it cannot be confidently used to make automatic decisions. Some of this information, such as fiscal or ethnic category, can even vary along an individual's life (See Wachtel[9]), so it was only used as supportive information when decisions had to be made in order to sort out inconsistent results.
- *Ayllus* should be considered decisive criteria when comparing grooms -or brides- against each other (in successive marriages). They should be more weakly used when comparing grooms against groom or bride's fathers. As the only *ayllus* cited in the source are the ones of the bride and the groom, when comparing an individual getting married with his/her reappearance as parent, the *ayllu* in the reappearance record is his/her son's or daughter's *ayllu*. But sometimes, *ayllus* are not properly inherited, as they are lost with the passing of time, or when an individual's child moves to an *hacienda*<sup>3</sup> <sup>4</sup>. So two cases are to be considered for identification: when there is *ayllu* coincidence, or when we lack *ayllu* information in one or both records. But if *ayllus* differ, the link is discarded.
- Women's names are sometimes not properly recorded, but, more important still, the use and transmission of surnames for women varies along the period <sup>5</sup>. Thus, special, more complex criteria should be used when comparing women from two different records. We began by creating a new field for brides: the *complete bride name* field, which is made up of the bride's name <sup>6</sup> concatenated with her mother and father's names. This way, we defined an equivalence function between women's names which decides that two women have the same name if the name of one of them has at least 2 terms included in the "complete name" of the other one, except when one of the terms is "MARIA", our most common first name <sup>7</sup>.

Based on this first experience with the *weight* function, and the conclusions drawn from it, a somewhat different approach was taken:

---

<sup>3</sup>Large agrarian estate.

<sup>4</sup>For more on *ayllu* inheritance, grouping into *moieties* and *ayllu* loss see Platt [6], Harris [3] or Izko [4].

<sup>5</sup>See Acosta and Tandeter [7].

<sup>6</sup>When we speak of *name*, we mean first name and surname together.

<sup>7</sup>55% of women in our corpus have "MARIA" as one of their first names.



For each linkable situation, we chose 1, 2 (or sometimes more) strong accepting criteria. So every identification made by these criteria was accepted, and then validated through the consistency checks described in the next section.

Some examples of acceptance criteria used can be seen in Table 1, and some compared results from each approach can be seen in the **Results** section.

Linkable pair	Criteria
Couple - Groom/Former Bride (strongest match criteria)	Same groom's first name and surname, groom's parents' names coincidence, no <i>ayllu</i> or <i>ayllu-moiety</i> contradiction, marital status in second record is "Widower", and distance between records is less than 50 years.
Couple - Groom/Former Bride (weaker match criteria)	Same first name and surname, no parents' names contradiction, no <i>ayllu</i> or <i>ayllu-moiety</i> contradiction, some coincidence between bride in first marriage and former bride in second one, marital status in second record is "Widower", and distance between records is less than 50 years.
Groom's parents - Groom's parents	Father's names coincidence, some element in Mother's names coincidence (complex criteria), and distance between records is less than 50 years.

Table 1: Examples of acceptance criteria for some linkable pairs.

## 5 Automatic Consistency Checks (Stage 4)

Once the record linkage process finishes, an underlying genealogy has been formed. But record linkage, which makes decisions based on pairs of records, can lead to inconsistencies. As stated by Wrigley [10], "[...] Perfect accuracy is beyond attainment in historical record linkage. [...]".

The decision was made to make consistency check a separate stage, so as to keep the record linkage process as simple as possible.

These types of possible inconsistencies were defined:

1. Individuals with more than one father or mother.
2. Individuals playing two different roles in the same record: sometimes, chains of identifications might lead to the case where, for example, we identify the groom and his father as being the same person. This has to be sorted out.
3. People married twice to the same person.

Cases where any of these problems arose were sought and displayed for a decision to be taken. The historian was provided with a tool for browsing through these cases and deciding which of the conflicting identifications (if any) was the correct one and which was a mistaken one. Some degree of arbitrariness is necessarily involved sometimes in this process. Figure 3 shows the tool used.

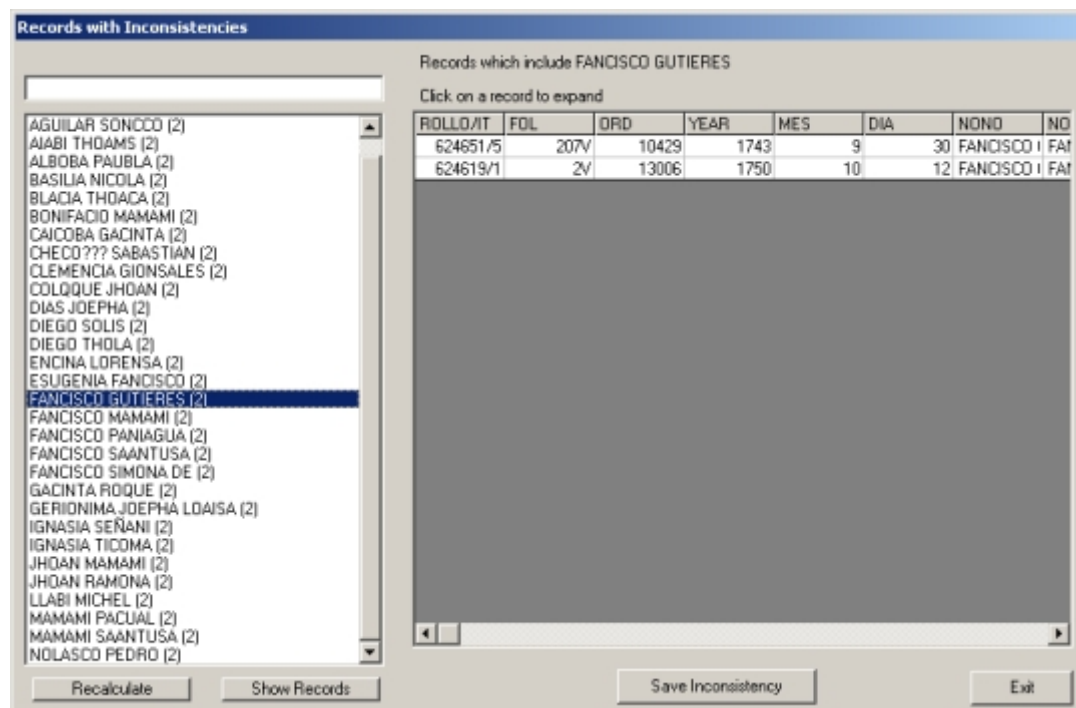


Figure 3: Software tool for checking and solving inconsistencies. The list on the left shows names involved in an inconsistency (in this case, people with two fathers). Clicking on a name, the inconsistent records are shown in full, so that a decision can be made.

This part of the process proved to be very painstaking, with many hours involved in inspecting individual cases.

## 6 Generation of GENOS format genealogy (Stage 5) and new inconsistency checks (Stage 6)

As mentioned before, GENOS is a software package, designed by Laurent S. Barry at the Social Anthropology Lab of the Collège de France in Paris, for the analysis of kinship.

It imports a genealogy file (with a specific format) and searches for certain kinds of marriage strategies in terms of consanguinity and affinity between the marrying parties.

We exported our genealogy, with the main inconsistencies solved, to GENOS format and sent it to Laurent Barry for processing.

During this process, certain new kinds of inconsistencies were detected. For example, a woman with more than 40 children was found, which is a clear identification

error, but undetectable by the inconsistency checks described in section 5. These new inconsistencies were again manually solved and a new genealogy was sent for processing through GENOS.

## 7 Results

In order to measure the success of this record linkage project, two basic metrics were selected:

1. Number of accepted links (after inconsistency checks).
2. Average genealogical depth obtained.

Our database of 11750 records is composed of 66739 individuals (meaning an average of 5.67 different people per record). Before the record linkage process, each of these individuals was assigned a unique identifying number. When two individuals were identified as being the same person, a new identifier was assigned to both appearances. After the process, 46265 different identifiers remained, meaning that 20474 links were accepted ( $66739 - 46265 = 20474$ ).

In order to measure genealogical depth, we started from the individuals without identified children, and traversed their ascendance tree in three different ways:

- By going up along the agnatic line (father).
- By going up along the uterine line (mother).
- By going up along any line (cognatic), trying to obtain the largest depth possible.

Figure 4 shows an example of a small genealogy, in order to demonstrate how genealogical depth is calculated.

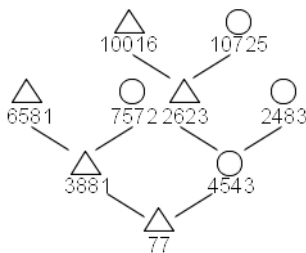


Figure 4: An example of a small genealogy.

In this example, the genealogical depth for individual 77 would be:

- Along the agnatic line, the depth is 3: 77 - 3881 - 6581.
- Along the uterine line, the depth is also 3: 77 - 4543 - 2483.
- Along the longest possible line (cognatic), we now get a depth of 4: 77, 4543 (mother), 2623 (father), 10725 (mother) or 10016 (father).

Table 2 shows the number (and percentage) of individuals with depths from 1 to 7 (our longest line, in the 120-year period, is reasonably no longer than 7 generations).

Metric	1	2	3	4	5	6	7
Males Agnatic	2581 (22.38%)	7471 (64.77%)	1249 (10.83%)	195 (1.69%)	38 (0.33%)	1 (0.01%)	0 (0.00%)
Females Agnatic	2660 (23.49%)	7198 (63.55%)	1237 (10.92%)	198 (1.75%)	31 (0.27%)	2 (0.02%)	0 (0.00%)
Males Uterine	2497 (21.65%)	7423 (64.35%)	1365 (11.83%)	228 (1.98%)	22 (0.19%)	0 (0.00%)	0 (0.00%)
Females Uterine	2557 (22.58%)	7296 (64.42%)	1243 (10.97%)	207 (1.83%)	21 (0.19%)	2 (0.02%)	0 (0.00%)
Males Cognatic	2456 (21.29%)	7228 (62.66%)	1254 (10.87%)	430 (3.73%)	123 (1.07%)	36 (0.31%)	8 (0.07%)
Females Cognatic	2527 (22.31%)	7048 (62.23%)	1185 (10.46%)	387 (3.42%)	135 (1.19%)	36 (0.32%)	8 (0.07%)

Table 2: Number and percentage of individuals with genealogical depths 1 to 7.

If we now compare these results to those of our first approach, based on the *weight* function, we get an idea of how this six stage approach proves itself much more stable and useful. The number of accepted links in the first approach depends on where we place the acceptance threshold, and how many of the links in the gray region are finally accepted. Anyhow, no reasonable accepting threshold accepts more than 500 or 600 links.

As for genealogical depth, we can observe in Table 3 that we have very few lines with depth greater than 3, and even in this case they don't exceed 3.5% of the total.

Metric	1	2	3	4	5	6	7
Males Agnatic	4923 (35.40%)	8531 (61.35%)	421 (3.03%)	30 (0.22%)	0 (0.00%)	0 (0.00%)	0 (0.00%)
Females Agnatic	4312 (30.69%)	9319 (66.34%)	395 (2.81%)	20 (0.14%)	1 (0.01%)	0 (0.00%)	0 (0.00%)
Males Uterine	4833 (34.76%)	8845 (63.61%)	225 (1.62%)	2 (0.01%)	0 (0.00%)	0 (0.00%)	0 (0.00%)
Females Uterine	4202 (29.91%)	9780 (69.62%)	65 (0.46%)	1 (0.01%)	0 (0.00%)	0 (0.00%)	0 (0.00%)
Males Cognatic	4786 (34.42%)	8582 (61.72%)	484 (3.48%)	53 (0.38%)	0 (0.00%)	0 (0.00%)	0 (0.00%)
Females Cognatic	4163 (29.63%)	9448 (67.26%)	407 (2.90%)	28 (0.20%)	2 (0.01%)	0 (0.00%)	0 (0.00%)

Table 3: Number and percentage of individuals with genealogical depths 1 to 7, using the weight function method.

## 8 Conclusions

1. The break-up of the record linkage process into these 6 stages, each of them self contained and testable on its own proved trustworthy and stable.
2. As stated by Adman et al. [1], we believe that the role of the historian is irreplaceable, being his judgment and contextual knowledge impossible to transfer to an automatic algorithmic process.
3. This set of tools, and these 6 stages, should suit most record linkage project requirements, for different kinds of sources, no matter the language, culture or time periods.

## 9 Future work

We believe a main task for the future is to generalize and put together these tools in a flexible and easy to use package, to be used by any record linkage project.

The weakest point at this moment is that we have no tool for describing a source structure, importing it from other digital media, and allowing the historian to define his own linkage rules in order to process it.

A tool of this kind can be developed, and it should suit a whole family of sources, basically comprising parish records (birth, marriage, death records), successive census or election polls.

## References

- [1] Peter Adman — Stephen W. Baskerville — Katharine F. Beedham. Computer-assisted record linkage: or how best to optimize links without generating errors. *History and Computing*, 4(1):2–15, 1992.
- [2] Gerrit Bloothoof. Corpus-based name standardization. *History and Computing - Special Issue on Record Linkage II*, 6(3):iii–vii, 1994.
- [3] Olivia Harris. De l’asymétrie au triangle. transformations symboliques au Nord de Potosí. *Annales E-S-C*, 33(5-6):1108–1125, Sept-Dec 1978.
- [4] Xavier Izko. *La doble frontera. Ecología, política y ritual en el altiplano central*. Hisbol/Ceres, La Paz, 1992.
- [5] R.J. Morris. Editorial - nominal record linkage: into the 1990s. *History and Computing*, 4(1), 1992.
- [6] Tristan Platt. Symétrie en miroir. le concept de yanantin chez les Macha de Bolivie. *Annales E-S-C*, 33(5-6):1081–1107, Sept-Dec 1978.
- [7] Luis Acosta — Enrique Tandeter. Fathers’ surnames transmission in two andean villages between 1692 and 1811. *Personal communication*, 2001.
- [8] Mario Boleda — Enrique Tandeter. *Démographie historique des Andes Centromeridionales*. Carlo Corsini (ed.).

- [9] Nathan Wachtel. *Le retour des ancêtres. Les Indiens Urus de Bolivie. XXe-XVIe siècle. Essai d'histoire régressive.* Paris, Gallimard, 1990.
- [10] E.A. Wrigley. *Identifying People in the Past.* Edward Arnold Publishers Ltd., 1973.