

Reconstitution par ordinateur de familles et de généalogies au service de la démographie historique : aspects méthodologiques et applications

Sergio DE IASIO et Enzo LUCCHETTI

Résumé

La généralisation du *micro-ordinateur personnel* a provoqué un regain d'intérêt pour la gestion automatisée des archives historiques, mais a aussi souligné la nécessité de recourir à des procédures intégrées et normalisées afin de parvenir à des résultats comparables. Un ensemble intégré de programmes est ici décrit, qui s'appuie sur le logiciel dBASE et assurent la gestion automatisée des actes de naissance (ou baptême), de mariage et de décès (ou permis d'inhumer), ainsi que la reconstitution des dossiers individuels, familiaux et généalogiques. Le logiciel a permis la reconstitution des liens de parenté à Civitella del Tronto (TE, Italie centrale) au départ des 74.241 actes de naissance (37.206), mariage (11.068) et décès (25.967) transcrits des registres civils, pour la période 1810-1973.

L'usage d'un logiciel intégré, normalisé et produisant des résultats comparables nous paraît susceptible de stimuler l'étude de populations d'effectif réduit car il atténue les contraintes inhérentes à leur traitement. L'analyse de la diversité et des constantes observées dans la démographie comparée des populations permet ainsi de préciser les phénomènes propres à chacune, à une période particulière ou à une étape de son développement historique, social ou économique. Enfin la constitution d'archives informatisées sur base de procédures communes favorise la comparaison et l'intégration des éléments issus de sources différentes : l'étude des comportements démographiques peut ainsi mieux bénéficier des contributions issues d'autres disciplines centrées sur l'étude de l'homme et des populations humaines.

Abstract

The diffusion of Personal Computer has increased the interest in the creation of computerized historical data archives. That has also pointed out a need of standard computer programs for data storage and processing, in order to get comparable results. In this work, we report an integrated package written in CLIPPER 5 (a popular compiler of dBASE III plus) for the data recording of individual acts (births, marriages and deaths) and the automatic family reconstitution. The routines were tested on the historical-demographic reconstruction of population of Civitella del Tronto (Abruzzo, Italy) from 1809 to 1973, referring to the births (37,206), marriages (11,068) and deaths (25,967) recorded in the Civil Status Registrations.

The increase of such studies could be the way to exceed the limits related to the small size of population. The analysis of changes of the demographic behaviours in different populations allows to underline the peculiar characteristics of each group or due to historical, social and economical events. Finally, the constitution of computerized archives, handled by common procedures, improves the comparison and the integration of components deriving from different sources: the study of demographic behaviours will be able to get benefit from contributions of other disciplines related to the study of man and human populations.

INTRODUCTION

Notre démarche poursuit la réalisation d'un double objectif : d'abord établir une méthode standardisée pour la gestion informatique des archives bio-démographiques et, ensuite, la rédaction d'un ensemble de logiciels conduisant à la reconstitution des populations.

Le premier objectif répond à un besoin vivement ressenti en Italie où la richesse des sources et des archives — exceptionnelles par leur ancienneté, leur abondance et leur qualité (Aymard et Delille, 1977) — souffre d'une fragmentation en éléments très dispersés. De ce fait, le patrimoine non seulement ne bénéficie que d'une mise en oeuvre malaisée et très partielle, mais se trouve aussi

menacé de disparition (Carucci et Schiavoni, 1989). Plusieurs chercheurs se sont consacrés à la réalisation de banques de données dans le souci, soit de garantir la conservation des sources historiques, soit, encore, de les rendre plus aisément accessibles. Cette démarche se trouve simplifiée, surtout pour les actes d'Etat Civil (naissance, mariage et décès) et les registres paroissiaux, en raison de la grande uniformité de structure, de contenu et de format de présentation qui les caractérise. Il s'avère ainsi utile de disposer de procédures pré-établies pour chaque type de source que tout chercheur puisse appliquer avec, bien entendu, les adaptations dont la nécessité s'imposerait (Schiavino, 1989).

Nous nous sommes assigné pour second objectif la programmation d'un logiciel, automatisant très largement au départ des données brutes, la restitution de la structure de la population et de son évolution historique. La réalisation de ce projet pourrait, selon nous, stimuler le développement des études consacrées à l'évolution des populations humaines et favoriser des collaborations plus larges et mieux articulées, par des échanges de fichiers sous un format commun et par la comparaison des résultats issus de méthodes similaires. La généralisation des micro-ordinateurs personnels a, en effet, ouvert de nouvelles perspectives très vastes pour cette orientation de la recherche où, cependant, un individualisme très marqué dans les techniques de récolte et de traitement des données mène à des résultats qui ne sont plus toujours comparables. Nous tenons, en particulier, à souligner que la multiplication des travaux se conformant à un même canevas et l'élargissement du débat qu'ils suscitent offrent, selon nous, le moyen plus approprié — et sans doute le seul à notre portée — pour surmonter les contraintes qu'impose l'étude de populations d'effectif réduit. L'analyse des variations et constantes observables dans la démographie comparée de plusieurs populations devrait permettre de discerner les phénomènes qui leur sont spécifiques de ceux qui marquent telle période ou encore telle étape de l'histoire de leur développement socio-économique (Lucchetti, 1978, 1989).

Certes, pour éviter le risque — déjà rencontré en démographie historique — de figer les hypothèses de travail dans un paradigme unique et pré-établi, et donc exclusivement répété (Schiaffino, 1989), il nous paraît souhaitable que le *logiciel* demeure très souple, que la structure de la banque des données, ainsi que son mode d'interrogation, se prêtent aisément aux adaptations requises par tel thème de recherche ou par les particularités d'une région ou d'une période déterminée.

Nous présentons ici un instrument qui se propose de remplir ces fonctions. Il s'agit d'un logiciel destiné à automatiser la gestion des actes de naissance, mariage et décès, et à les structurer de manière à réaliser la reconstitution des dossiers individuels, familiaux et généalogiques. Par ces programmes — qui fonctionnent sous dBASE — nous avons traité les données communiquées par l'Etat Civil de Civitella del Tronto (Teramo, Italie centrale) : soit au total 74.241 actes dont 37.206

naissances, 11.068 mariages et 25.967 décès transcrits de registres qui couvrent la période qui s'étend de 1810 à 1973 (De Iasio *et al.*, 1989).

SAISIE DES DONNEES

Le choix des variables et la structuration des fichiers — première étape de la démarche — déterminent de manière irréversible les possibilités de traitement ultérieur. D'autres exigences s'imposent : le *logiciel* — et l'*architecture matérielle* sur laquelle il s'appuie — devraient bénéficier d'une ample diffusion et d'un usage très souple. En effet, ils doivent enregistrer les informations sous un format très proche de celui de la source originale et favoriser un traitement rapide et souple, plus particulièrement la détection et la gestion de relations entre champs de la base (*colonnes* ou *variables*) et entre les actes (*records*). Le produit doit enfin s'adapter à divers cas particuliers et pouvoir fonctionner non seulement sur les puissants ordinateurs (*mainframes*) accessibles dans les centres de calcul, mais aussi sur des micro-ordinateurs portables jusque dans les locaux des archives.

Les données issues des actes de naissance, mariage et décès sont chargées, respectivement, dans trois *fichiers* distincts, qui sont alimentés et consultés séparément ou non (par exemple, celui des mariages pour des analyses spécifiques sur le choix du conjoint). L'opérateur est aidé par un masque de saisie qui lui rappelle quelles sont les informations à introduire : il lui suffit de suivre la grille proposée et d'encoder les données dans l'ordre prévu et selon les indications affichées dans la partie inférieure de l'écran. La figure 1 illustre les trois masques de saisie tels qu'ils apparaissent à l'opérateur. Le système, nous le soulignons, est d'un usage très souple : on peut modifier très facilement le schéma d'affichage (ajouter ou éliminer de l'écran une zone d'enregistrement), ou l'ordre d'introduction des données, pour mieux se conformer à la structure des actes originaux. L'opérateur bénéficie d'une *fonction d'aide* qui lui rappelle quelques conventions — par exemple, comment formuler les âges en jours ou en mois — conventions qu'on peut aménager très facilement. Enfin, il existe un *fichier de notes* — maintenu automatiquement en relation avec le *fichier* principal par la numérotation des *enregistrements* (*records*) — où l'utilisateur place des remarques particulières, présentes sporadiquement, qu'il juge néanmoins significatives.

BIRTH EVENT # / []		Progr.Numb.Code []	
Birth Recording Date []-[]-[]		Birth Place []	
Actual Birth Date []-[]-[]		Residence Place []	
BORN	Surname []		
	1 st Name []		
	2 nd Name []		
	3 rd Name []		
	Sex [] (M/F)		
FATHER	1 st Name []		
	2 nd Name []		
	Age []	Marital St. []	Profession []
MOTHER	Surn. []		
	1 st Name []		
	2 nd Name []		
	Age []	Marital St. []	Profession []
Annotations []			

MARRIAGE EVENT # / []		Progr.Numb.Code []	
Recording Date []-[]-[]		Marriage Date []-[]-[]	
		Place []	
H	Surname []	W	Surname []
U	1 st Name []	I	1 st Name []
S	2 nd Name []	F	2 nd Name []
B	Birth date []-[]-[]	E	Birth date []-[]-[]
	Age []		Age []
A	Birth place []		Birth place []
N	Mar.St. []		Mar.St. []
	Resid. []		Resid. []
D	Profess. []		Profess. []
FATHER	1 st Name []	FATHER	1 st Name []
	2 nd Name []		2 nd Name []
	Age []		Age []
	M.S. []		M.S. []
	Prof. []		Prof. []
MOTHER	Surname []	MOTHER	Surname []
	1 st Name []		1 st Name []
	2 nd Name []		2 nd Name []
	Age []		Age []
	M.S. []		M.S. []
	Prof. []		Prof. []

DEATH EVENT # / []		Progr.Numb.Code []	
Recording Date []-[]-[]		Death Date []-[]-[]	
		Place []	
D	Surname []		
E	1 st Name []		
A	2 nd Name []		
D	3 rd Name []		
	Sex [] (M/F)	Age []	Birth date []-[]-[]
	Residence place []	Marital Status []	
		Profession []	
FATHER	1 st Name []		
	2 nd Name []		
MOTHER	Surname []		
	1 st Name []		
	2 nd Name []		
CONSORT	Surname []		
	1 st Name []		
	2 nd Name []		
Annotations []			

Figure 1 : Grilles pour l'entrée des données (actes de naissance, mariage et décès).

Les données chronologiques sont saisies selon le *format de données* du dBASE : le programme contrôle celles-ci, soit pour leur cohérence interne, soit pour leur compatibilité avec d'autres données (ou âges) figurant sur le même acte. Les informations de type caractère (noms, prénoms, localités, professions) sont enregistrées sans autre codage, telles qu'elles apparaissent dans le document original, sans aucune normalisation ni codification a priori. Un même nom ou prénom peut connaître diverses versions orthographiques : la normalisation de ces variantes constitue un problème que l'opérateur doit résoudre au cas par cas. Le logiciel lui affiche en permanence, pour chaque type d'information alphabétique, la liste des formes déjà connues et leurs fréquences; au moyen d'une table de conversion fournie par le programme, le chercheur peut ramener toutes les variantes à la seule forme convenue.

Le codage des informations de type caractère — nécessaire pour gagner de la place en mémoire centrale et pour obtenir un traitement plus rapide — s'effectue de manière complètement automatique après la saisie des données, au moyen d'un *utilitaire* qui les range par ordre alphabétique et assigne aux variantes un code commun. La normalisation et codage des données ne jouent qu'un rôle fonctionnel : la version initiale se trouve toujours conservée dans la forme issue de la source originale.

RECONSTITUTION D'UNE POPULATION

La reconstitution historico-démographique d'une population s'effectue en deux étapes : la première construit les dossiers individuels, c'est-à-dire la réunion des actes relatifs à un même individu; la seconde opère la reconstitution des familles et des généalogies, c'est-à-dire l'identification des liens de parenté et de filiation.

La méthode de reconstitution au départ des actes d'Etat Civil et registres paroissiaux (naissance ou baptême, mariage, décès ou permis d'inhumer) suit toujours le même principe et n'altère en rien les données originales car la relation qui unit l'ensemble des renseignements propres à une même entité (personne ou famille) se construit sur base de codes qui renvoient vers le numéro identificateur de chaque acte (*enregistrement* ou *record*).

Le logiciel peut être utilisé soit d'une façon

complètement automatique, soit en mode interactif qui permet à l'opérateur d'intervenir à tout moment qu'il juge opportun.

RECONSTITUTION DES DOSSIERS INDIVIDUELS

La première étape concerne la constitution des dossiers individuels, c'est-à-dire la réunion des actes relatifs à un même individu. En pratique, le programme de reconstitution établit d'abord les liaisons entre les actes de décès et de naissance, auxquelles viennent ensuite se joindre les actes de mariage.

Voici les indications qui guident l'identification des actes propres au même individu : le nom, le premier et le second prénom du sujet, le sexe, les dates des événements, l'âge, le prénom du père, le nom et le prénom de la mère, toutes informations presque toujours consignées dans les documents de l'Etat Civil. Généralement l'inventaire des actes qui lui sont communs s'opère sur base d'une clé d'accès formée par le code du nom et du prénom du sujet. Le système offre, de plus, la faculté de modifier la clé d'accès et de prendre, ainsi, en considération les variantes éventuelles des prénoms ou du nom. Afin d'optimiser et d'accélérer le processus, le chercheur fixe un délai maximum au programme pour établir les liaisons. Le traitement va s'effectuer en plusieurs étapes : d'abord le programme détecte et mémorise les liaisons qui réunissent l'ensemble des informations attendues; ensuite les conditions requises pour établir le lien deviennent moins restrictives et celui-ci sera accepté malgré l'absence ou la discordance de l'un ou de l'autre élément. De toute manière, le chercheur peut fixer lui-même les règles, l'ordre et la rigueur des comparaisons. Enfin, à chaque pas, il peut intervenir sur toutes les liaisons proposées, les réexaminer, résoudre les cas douteux, et modifier les solutions proposées par le logiciel.

RECONSTITUTION DES FAMILLES ET DES GENEALOGIES

L'algorithme de reconstitution de familles applique une méthode analogue, très simple. Au moyen d'une clé formée par le nom et le prénom du conjoint, il identifie l'ensemble des titulaires d'actes issu d'un couple commun. L'opération se trouvera accélérée si la date de naissance (ou l'âge) de la mère est connue et permet ainsi de déterminer la période durant laquelle elle est capable de

procréer et à laquelle sera limitée la recherche d'une éventuelle progéniture (la date du mariage ne fixe pas la limite inférieure car elle écarterait les naissances prénuptiales). L'identification des liens de parenté et de filiation s'effectue, elle aussi, selon une logique dont la rigueur s'exerce par paliers décroissants et dont l'opérateur peut choisir les modalités et le niveau d'exigence.

La structure finale ainsi obtenue rappelle la fiche familiale proposée par Fleury et Henry (1976) (les reconstitutions, pour être comparables, doivent se ramener à ce schéma); elle apparaît sur l'écran telle qu'à la figure 2. Dans la partie supérieure se trouvent les informations relatives au couple (date de mariage, date de la fin de l'union, sa durée, le nombre d'enfants, puis la date de naissance, les parents des conjoints, l'âge au mariage, la date et l'âge au décès de chacun des époux). Les données relatives aux enfants occupent la moitié inférieure, présentées selon le rang de naissance : sexe, prénom, date de naissance, date de décès et âge au décès, date(s) de mariage(s), âge au ma-

riage et nom du conjoint. Pour chaque descendant sont repris, à gauche, l'âge de la mère à sa naissance, le temps écoulé, exprimé en mois, depuis la date du mariage ou la naissance antérieure.

D'autres analyses peuvent être générées très aisément — par des *utilitaires* — au départ de la fiche familiale : par exemple le calcul de l'âge moyen du père et de la mère à la naissance des enfants, la durée de l'intervalle protogénésique, la durée moyenne des intervalles intergénéraliques et, par tranche de cinq ans durant la période de sa vie où l'épouse peut procréer, le nombre d'années de mariage et le nombre d'enfants.

La reconstitution des généalogies, par la voie ascendante ou descendante, s'obtient aisément sur base des familles déjà reconstituées : la fiche familiale permet de connaître les enfants issus du mariage et d'explorer sa descendance ou, au contraire, de remonter vers les ascendants de génération en génération; les indications relatives aux frères et soeurs permettent aussi de parcourir toutes les branches collatérales.

Marr.Date 1909.11.25		BERRETTONI PACIFICO		CROCETTI ROSA					
Union End 1946.12.27		b.1885.03.19 Age Marr.24		b.1884.09.06 Age Marr.25					
Marriage Duration: years 37;months 1		d.1966.05.15 Age d.81		d.1946.12.27 Age d.62					
Offspr:11 M 7 F 4		Parents : ALETO		Parents : FABIO					
Spouse:5 M 3 F 2		DIANETTI FILOMENA		MARTELLI MARGHERITA					
Age	Durat.	C H I L D		B I R T H		D E A T H		M A R R I A G E	
F	Months	Sex	First name	Date	Date	Age	Date	Age	SPOUSE
28	37	M	MARCELLO	1913.01.02			1937.04.17	24	ANNA
29	50	F	FILOMENA	1914.01.28	1949.04.07	35	1935.05.25	21	CESARINO
30	68	M	ADOLFO	1915.07.09	1963.02.07	47			
32	37	M	ALFONSO	1917.01.31	1917.04.19	m 2			
33	101	F	PALMA	1918.03.24					
34	115	M	ADAMO	1919.05.09			1948.01.05	28	PASQUALINA
36	135	M	GIOVANNI	1921.01.01	1922.02.28	1			
37	151	F	EUROSIA	1922.04.30			1942.09.03	20	GIUSEPPE
40	184	M	ALETO	1925.01.17	1925.01.31	d14			
40	184	F	ANTONIA	1925.01.17	1925.01.17	d 0			
42	210	M	DOMENICO	1927.03.06			1956.01.19	28	SALUTARIA

Figure 2 : Exemple de la fiche de famille obtenue par la reconstitution.

CONCLUSIONS ET PERSPECTIVES

Nous ne soulignerons plus ici l'intérêt que présente la constitution d'archives informatisées et centralisées dans le but, non seulement d'inventorier et de cataloguer la documentation historique, mais aussi de combler ses lacunes, de la conserver et surtout d'en rendre la consultation plus facile. La centralisation de plusieurs archives pourrait permettre l'étude comparée des sources : on évaluerait la fidélité et le caractère exhaustif des informations; des sources différentes pourraient être confrontées et complétées; des corrélations nouvelles pourraient faire l'objet d'études et engendrer de nouvelles connaissances.

La création d'un logiciel intégré reconstituant des populations du passé répond, elle aussi, à un besoin ressenti depuis longtemps.

Dans ces deux types d'application, la diffusion croissante des *micro-ordinateurs personnels*, doués d'une large capacité de mémoire et d'une vitesse de traitement satisfaisante, a créé de nouvelles potentialités suscitant un regain d'intérêt pour ces travaux. Surtout fut perçue l'opportunité de développer une approche multifactorielle, déterminante pour la compréhension des phénomènes démographiques. Non seulement les divers événements enregistrés à l'Etat Civil et marquant la biographie d'un individu doivent être mis en relation, mais aussi associés à ceux de sa parentèle : enfants, parents, frères, soeurs, etc... (Skolnick *et al.*, 1978). Au-delà, les phénomènes démographiques doivent se replacer dans le contexte social et historique où ils sont vécus (Soliani, 1989). Il suffit, par exemple, de rappeler que les deux composantes du comportement matrimonial — choix du conjoint et comportement reproducteur — s'influencent mutuellement au sein d'un processus rétroactif : la première incarne l'aboutissement d'un jeu d'interactions issues de facteurs très divers (religion, coutumes, traditions, conditions sociales et économiques de la famille et du groupe, etc...), et la seconde constitue le mécanisme par lequel la population perpétue, ou modifie au cours du temps, sa structure, sa composition, son patrimoine culturel, social, biologique et génétique (Lucchetti, 1989).

L'usage d'une méthode normalisée, d'un emploi facile, susceptible d'une large diffusion et capable de fournir des résultats aisément compara-

bles peut stimuler le développement d'une telle approche.

Remerciements

Les auteurs remercient Mr. Jean-Louis Slachmuylder pour la révision du texte français.

Bibliographie

- AYMARD, M. et DELILLE, G., 1977. La démographie historique en Italie : une discipline en mutation. *Annales de Démographie historique* : 447-461.
- BOETSCH, G. et SEVIN, A., 1990. Application de la méthode généalogique à l'étude d'une population ouverte (Chateauponsac-Haute-Vienne). *Approche pluridisciplinaire des isolats humains*. Paris, I.N.E.D. : 384-400.
- CARUCCI, P. et SCHIAVONI, C., 1989. Le scritture parrocchiali in Italia tra XVI e XX secolo. Un patrimonio in bilico tra Stato e Chiesa, in cerca di tutela e ordinamento. In : G.Coppola et C.Grandi (éd.) : *La «conta delle anime» Popolazioni e registri parrocchiali: questioni di metodo ed esperienze*. Bologna, Soc. Editrice Il Mulino : 21-34.
- DE IASIO, S., GHISOLFI, G. et LUCCHETTI, E., 1989. Analisi microdemografiche sulla popolazione di Civitella del Tronto. La ricostruzione delle storie individuali (1810-1973). In : S.I.D.E.S. (éd.) : *Utilizzazione del Personal computer nelle ricerche di Demografia storica*. *Bollettino di Demografia storica*, 9 : 77-88.
- FLEURY, M. et HENRY, L., 1976. *Nouveau manuel de dépouillement et d'exploitation de l'état civil ancien* (2e édition). Paris, I.N.E.D.
- LUCCHETTI, E., 1978. Aspetti genetici e demografici nello studio di popolazioni umane. *Antropologia contemporanea*, 1 : 41-57.
- LUCCHETTI, E., 1989. Integrazione tra informazioni in analisi su popolazioni umane. In : S.I.D.E.S. (éd.) : *Utilizzazione del Personal computer nelle ricerche di Demografia storica*. *Bollettino di Demografia storica*, 9 : 40-50.
- SCHIAFFINO, A., 1989. Presentazione. In : S.I.D.E.S. (éd.) : *Utilizzazione del Personal computer nelle ricerche di Demografia storica*. *Bollettino di Demografia storica*, 9 : 4-6.

SKOLNICK, M., BEAUCHAMP, P., CHARBONNEAU, H.,
LEGARE, J., CORSINI, C. et LE MEE, R., 1978.
Colloque sur les méthodes de reconstitution
automatique des familles. *Soc. de Dém. hist.*
(dh) Bulletin d'information, 25.

SOLIANI, L., 1989. Per un rilancio delle monografie
parrocchiali e comunali. *In* : S.I.D.E.S. (éd.) :
Utilizzazione del Personal computer nelle ri-
cerche di Demografia storica. *Bollettino di De-*
mografia storica, 9 : 134-143.

Adresse des auteurs : S. DE IASO et E. LUCCHETTI
Département de Biologie évolutive
Université de Parme
Viale delle Scienze
I-43100 Parme (Italie)

Manuscrit reçu le 31 juillet 1992