

La reconstitution des familles par Dbase

Autor(en): **Sorgesa Miéville, Beatrice**

Objektyp: **Article**

Zeitschrift: **Geschichte und Informatik = Histoire et informatique**

Band (Jahr): **5-6 (1994-1995)**

PDF erstellt am: **08.08.2024**

Persistenter Link: <https://doi.org/10.5169/seals-118837>

Nutzungsbedingungen

Die ETH-Bibliothek ist Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Inhalten der Zeitschriften. Die Rechte liegen in der Regel bei den Herausgebern.

Die auf der Plattform e-periodica veröffentlichten Dokumente stehen für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden.

Das Veröffentlichen von Bildern in Print- und Online-Publikationen ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Die systematische Speicherung von Teilen des elektronischen Angebots auf anderen Servern bedarf ebenfalls des schriftlichen Einverständnisses der Rechteinhaber.

Haftungsausschluss

Alle Angaben erfolgen ohne Gewähr für Vollständigkeit oder Richtigkeit. Es wird keine Haftung übernommen für Schäden durch die Verwendung von Informationen aus diesem Online-Angebot oder durch das Fehlen von Informationen. Dies gilt auch für Inhalte Dritter, die über dieses Angebot zugänglich sind.

La reconstitution des familles par Dbase

Beatrice Sorgesa Miéville

L'adaptation informatique de la méthode développée par Henry et Fleury¹, mise au point dans le cadre d'une recherche consacrée à l'impact de l'industrialisation horlogère sur les comportements et les structures familiaux², comprend environ 80 programmes qui permettent, à partir de 4 fichiers de base – Naissances, Décès, Epoux et Epouse –, d'aboutir d'une part à la reconstitution semi-automatique des familles; d'autre part à la réalisation complètement automatique des calculs destinés à l'étude de la nuptialité et de la fécondité (âge au mariage, durée de l'union, âge de la mère à la maternité, intervalle entre les naissances etc.).

Travail de néophyte, du point de vue informatique, il mériterait de nombreuses mises à jour et corrections, d'autant plus qu'il a été conçu sur la base des possibilités offertes par Dbase II, possibilités bien plus limitées que celles de Dbase IV, notamment en ce qui concerne le nombre maximal de champs de chaque fichier.

Malgré ses imperfections, le matériel élaboré pourrait constituer une base de départ pour des chercheurs confrontés au problème de la reconstitution des familles et ne disposant pas d'une infrastructure informatique lourde et/ou n'ayant pas une formation informatique. La souplesse, la vaste gamme d'applications possibles et surtout la facilité d'utilisation du logiciel Dbase en font à coup sûr un instrument précieux pour l'historien.

La démarche qui n'exclut pas l'intervention manuelle du chercheur lors de certaines phases – intervention qui devient d'autant plus fréquente et astreignante que la documentation de base (les registres d'état civil) est lacunaire et peu formalisée – s'articule en 7 phases dont les buts sont les suivants:

Première phase: après la saisie des données – actes de naissance/baptême, de décès et de mariage – il s'agit de rassembler, dans la mesure du possible, toutes les informations disponibles concernant le même individu dans une seule et même fiche: l'acte de mariage, à défaut l'acte de décès.

Cela revient, du moins pour les individus dont l'acte de mariage a été relevé, à remplir la partie supérieure des fiches de familles de 1^o type, selon la terminologie Henry-Fleury.

Concrètement il s'agit d'opérer une série de couplages (mariages/naissances; mariages/décès; décès/naissances etc.) dans un ordre préétabli que

1 H. Fleury et L. Henry, *Nouveau manuel de dépouillement et d'exploitation de l'état civil ancien*, Paris, INED, 1965.

2 B. Sorgesa Miéville, *De la Société traditionnelle à l'ère industrielle : les comportements familiaux face au changement économique. Mutations démographiques d'un village horloger du Jura Neuchâtelois (Fleurier, 1727-1914)*, Neuchâtel, Société d'histoire et d'archéologie, 1992.

l'ordinateur peut faciliter et accélérer mais rarement effectuer de manière automatique en raison de la forte homonymie, de la variabilité des prénoms le long du cycle de vie individuel et de l'orthographe des noms de famille encore mal fixée qui caractérisent la période ancienne. Des performances meilleures sont possibles lorsque la documentation de base est bien formalisée (pour Neuchâtel dès le 2^e quart du XIX^e siècle).

La *deuxième phase* consiste à rechercher les enfants nés des couples dont le mariage a été relevé ou éventuellement retrouvé par remariage. Cela revient, selon le jargon de la démographie historique, à remplir la partie inférieure des fiches de famille de 1^o type. Concrètement il s'agit d'opérer une deuxième série de couplage, toujours dans un ordre préétabli (naissances/mariages; décès/mariage; mariage/mariage) que l'ordinateur peut, dans une certaine mesure, effectuer de manière automatique grâce à des programmes fondés sur des tests d'identification. Cela n'évite pas un certain nombre de contrôles manuels au chercheur, mais encore une fois le travail est simplifié et accéléré par l'utilisation de l'informatique.

La *troisième phase* consiste à rassembler tous les enfants des couples dont le mariage n'a pas été ni relevé ni retrouvé, ce qui, dans la terminologie Henry-Fleury, revient à remplir la partie inférieure des fiches de famille de 2^o type. Concrètement il s'agit de trier les fichiers résiduels (c'est-à-dire naissances, mariages et décès non encore transcrits en enfants) et de regrouper les enregistrements par famille (enfants et parents).

Dans cette troisième phase également l'ordinateur a exécuté automatiquement une partie du travail seulement et n'a pas pu éviter des contrôles manuels réguliers.

Le but de la *quatrième phase* est celui de procéder à une première vérification destinée à résoudre un certain nombre de problèmes laissés en suspens lors des trois premières phases tels que «oublis» d'enfants suite à l'utilisation de critères de sélection trop stricts et mal adaptés à la qualité des sources ou multiplication erronée des enregistrements concernant un même enfant ou un même couple suite à des couplages manqués entre les actes ou encore création d'une famille de 2^o type à partir d'une naissance illégitime. Cette vérification a été effectuée par des contrôles exclusivement manuels, que l'ordinateur a toutefois simplifiés et accélérés.

La *cinquième phase* a pour but de rassembler toutes les informations disponibles concernant les conjoints des couples dont le mariage a été retrouvé à travers la naissance des enfants, ce qui revient à remplir la partie supérieure des fiches de famille de 2^o type selon la terminologie Henry-Fleury.

Concrètement il s'agit de procéder à une troisième série de couplages selon un ordre préétabli (décès/mariages, naissances/mariages) qui pour l'essentiel ont pu être accomplis de manière automatique par l'ordinateur

grâce à une série de programmes de transcription fondés sur des tests d'identification.

Théoriquement à la fin de cette cinquième phase les fiches de toutes les familles dont la naissance d'un enfant au moins a été relevée sont, dans la mesure du possible, complétées par toutes les informations disponibles concernant tant les parents que les enfants. En réalité une partie de l'information disponible n'a pas été complètement exploitée en raison de critères de couplage trop stricts et mal adaptés à la documentation de base, qui se prête mal parfois à une formalisation trop poussée. De plus le risque d'un doublement illicite d'enregistrements concernant un même enfant ou un même couple subsiste malgré la vérification de la 4^o phase. Enfin des faux couplages ont pu échapper à la vigilance du chercheur.

La *sixième phase* a donc pour but d'exploiter au mieux les actes résiduels et de contrôler et corriger les erreurs commises au cours des cinq premières phases.

Ces opérations ont été pour l'essentiel accomplies manuellement, mais l'aide de l'ordinateur a été essentiel sur deux points :

1. L'élimination des enfants inscrits plus d'une fois dans la même fiche de famille
2. Le contrôle de la validité de la reconstitution grâce à des programmes qui testent automatiquement la vraisemblance et la cohérence de chaque histoire familiale. Sur la base d'une série de critères naturels et sociaux tels que deux enfants nés à la même date ne portent généralement pas le même prénom; l'intervalle entre naissances n'est généralement pas inférieur à 9 mois; on ne peut pas avoir un enfant légitime avant le mariage; on ne peut pas mourir avant d'être né; on accouche en général avant 50 ans etc. Chaque cas douteux sur un ou plusieurs aspects est signalé au chercheur et soumis à son jugement.

La *septième phase* est complètement automatique. Il s'agit de calculer pour chaque famille l'ensemble des paramètres démographiques – base des différentes statistiques qui fondent l'analyse de la nuptialité, de la natalité et la fécondité – tels que l'âge au mariage, l'âge au décès, l'âge en fin d'observation, la durée de l'union, l'âge de la mère à la maternité, l'intervalle séparant les naissances, le nombre de naissances intervenues dans chaque groupe d'âge de la mère, le nombre de femmes-années passées par chaque femme dans chaque groupe d'âge etc.

A la fin de cette septième phase le chercheur dispose désormais de 5 fichiers soit : 1. *Epoux* et 2. *Epouses* contenant les informations relatives à chaque conjoint; 3. *Enfant* contenant les informations relatives aux enfants de chaque couple; 4. *Mariage* contenant les informations relatives à l'histoire

conjugale du couple; 5. *Fécondité* contenant les informations relatives à l'activité de procréation de chaque couple. Tous les enregistrements concernant une même famille sont reliés entre eux par un même numéro d'identification.

Selon l'analyse démographique que l'on désire effectuer, l'ordinateur est en mesure d'établir à partir des champs sélectionnés et en principe en faisant appel simplement aux fonctions intégrées de Dbase, les statistiques qui serviront à l'étude des caractéristiques de la natalité, de la fécondité, de la nuptialité et éventuellement de la mortalité.

L'apport de Dbase dans cette démarche peut être résumé ainsi:

1. Fonctions intégrées Dbase et notamment fonction INDEX qui permet d'établir rapidement des listes en ordre alphabétique ou chronologique, triées éventuellement selon certains critères (lieu de naissance, état civil au mariage etc.) ce qui facilite et accélère les couplages et les contrôles manuels (phase 1, phase 2, phase 3, phase 4, phase 5, phase 6)
2. Programmes de transcription qui permettent de transcrire automatiquement d'un acte à l'autre les informations sélectionnées et «marquent» les actes transcrits (phase 1, phase 6)
3. Programmes de couplage, transcription et marquage automatique fondés sur des tests d'identification sur la base de critères tels que nom et prénom de l'individu considéré, noms et prénoms des parents, date de naissance etc. (phase 1, phase 2, phase 3, phase 5)
4. Programmes de vérification de vraisemblance (phase 6)
5. Programmes de calcul automatique (phase 7).

Les «marquages» réguliers se révèlent très utiles parce qu'ils permettent, à la fin de chaque phase, de scinder les fichiers utilisés pour les couplages en les épurant progressivement des enregistrements «inutiles», c'est-à-dire déjà transcrits. En réduisant au fur et à mesure l'étendue des fichiers de base, cela permet de mieux cibler les contrôles manuels et réaliser ainsi un gain de temps parfois considérable.

Sans complètement libérer le chercheur de la fastidieuse corvée que représente la reconstitution des familles, le logiciel Dbase la rend moins lourde et bien plus rapide et ouvre des nouvelles perspectives d'exploitation non seulement en histoire démographique, mais également en histoire sociale (ex. reconstitution de réseaux de parenté, étude de la mobilité sociale sur plusieurs générations). Son avantage principal réside dans sa facilité d'accès, même pour le néophyte.

Le succès de l'entreprise reste toutefois lié à une condition essentielle: une connaissance approfondie des sources qui seule permet la réalisation de programmes efficaces et adaptés à la documentation utilisée. Et en cela l'informaticien ne peut pas remplacer l'historien.