

I. Scollar, B. Weidner, I. Herzog
**Ein portables Seriations-Programmpaket mit dynamischer
Speicherplatzverwaltung in PASCAL**

Es wurden zwölf Programme in PASCAL geschrieben, um eine Seriation in annähernd linearer Zeit durchzuführen, wobei ein neuer Algorithmus von P. Ihm verwendet wird. Zur Zeit arbeiten die Programme auf Besetzungsmatrizen, deren Größe nur durch den verfügbaren Speicherplatz begrenzt ist. Eine Erweiterung zur Bearbeitung von Häufigkeitsmatrizen ist geplant.

Das Seriations-Paket enthält Programme zur interaktiven Dateneingabe (Grab-orientiert, Typ-orientiert oder gemischt) und zur Überprüfung, wie gut die Daten verknüpft sind, um die Seriationsergebnisse zu drucken und die Resultate verschiedener Seriationsläufe oder Eingabereihenfolgen zu vergleichen. Testdaten verschiedenen Umfangs sowie ein kleines Benutzer-, Installations- und Programmhandbuch werden in maschinenlesbarer Form mitgeliefert. Standardisierte Datenformate erlauben es, Zwischenergebnisse zu editieren und Programme wiederholt zum Korrigieren der Daten zu benutzen, wenn dies nötig sein sollte.

Die Programme werden nur als maschinenlesbare Quellprogramme weitergegeben. Alle systemabhängigen Programmteile sind in separaten Dateien gespeichert. Diese können nach Bedarf geändert und mitkompiliert werden. Bei allen Programmen wird die Bildschirm- und -ausgabe von jeweils einer zentralen Kontrollroutine (message handler) ausgeführt. Diese können modifiziert werden, um sie der Sprache, die der Benutzer wünscht, oder dem lokalen System anzupassen. Alle Konstanten, die den Speicherplatzverbrauch beeinflussen, befinden sich in einer eigenen Datei, die je nach Speicherkapazität Ihres Computers editiert werden kann und bei Kompilation miteingefügt wird. Bereits geringe Kenntnisse in PASCAL werden in der Regel ausreichen, dieses Programmpaket auf einer anderen Maschine zu implementieren. In allen Programmen wird der Speicherplatz dynamisch bereitgestellt bis zu den Grenzen, die abhängig vom jeweiligen Computer bei Kompilation festgelegt werden.

Der Algorithmus von Ihm, der im Hauptseriationsprogramm benutzt wird, führt mit schneller Konvergenz zu einer eindeutigen Lösung des Seriationsproblems. Der Algorithmus ist äquivalent zur Benutzung der Korrespondenzanalyse auf den Daten mit Maximierung des Produktmomentkorrelationskoeffizienten. Dies ist die schnellste bekannte statistisch stabile Methode. Das Programm bietet eine Auswahl von Stopbedingungen an, wobei die Iterationen fortgeführt werden bis entweder eine Fehlernorm die gegebene Fehlerschranke unterschreitet oder bis eine stabile Typ- und Grabreihenfolge erreicht ist. Letzteres wird durch ein schnelles Sortierverfahren in der Endphase des Programms überprüft. Die Rechenzeit für einen größeren Datensatz (720 Gräber, 225 Typen mit 2000 Objekten) liegt mit einem großen Minicomputer in der Größenordnung von Minutenbruchteilen. Daten und Programm brauchten dabei 64 K Bytes Speicherplatz. Die Programme laufen auf jedem Computer mit einem gewöhnlichen PASCAL-Kompiler sowie genug Speicherplatz, um ausführbaren Code und Daten zu speichern, also von 48 K Bytes aufwärts. Gleitkommahardware ist nützlich, aber nicht Voraussetzung.

Das Dateneingabeprogramm bietet die Auswahl zwischen einem Typ-in-Grab und einem Grab-beinhaltet-Typ-Modus. Dieses Programm überprüft die Syntax der Eingabedaten sofort und meldet auftretende Fehler. Die Dateneingabe kann unterbrochen und zu einem anderen Zeitpunkt fortgesetzt werden. Damit kann man Daten hinzufügen und einen anderen Eingabemodus benutzen. Ein Sortier- und Mischprogramm bringt die Daten in die standardisierte Form, die als Eingabeformat für Folgeprogramme vorgeschrieben ist. Ein Testprogramm überprüft Dateien, die durch eine andere Eingabemethode erzeugt wurden, wie etwa durch ein Datenkonversionsprogramm, das bereits vorhandene Datensätze umwandelt.

Ein weiteres Prüfprogramm analysiert mit Hilfe mengentheoretischer Methoden, ob die Gräber miteinander verknüpft sind und findet voneinander isolierte Gräber und Grabgruppen, die in verschiedene Ausgabedateien geschrieben werden können. Das Programm ist konzeptionell der Arbeit mit Venn-Diagrammen äquivalent.

Das Druckprogramm kann auch sehr große Matrizen auf nebeneinander zu legende Papierbahnen drucken. Dabei erscheinen das Seriationsergebnis oder die Eingabereihenfolge als Besetzungsmatrix.

trix, wobei die Grab- bzw. Typnamen entlang der zugehörigen Achsen gedruckt werden.

Ein Vergleichsprogramm für zwei verschiedene Seriationsergebnisse oder für ein Seriationsergebnis und die Eingabereihenfolge druckt ein Vergleichsdiagramm (Scatterplot), das auch kleine Abweichungen von perfekter Übereinstimmung sichtbar macht. Außerdem errechnet das Programm den Korrelationskoeffizienten zwischen den zu vergleichenden Reihenfolgen.

Folgende Erweiterungen und zusätzlichen Programme sind geplant: die oben schon erwähnte Erweiterung auf Häufigkeitsmatrizen, mit entsprechenden Dateneingabe- und Prüfprogrammen; Reduktion von Häufigkeits- auf Besetzungsmatrizen; eine Schnittstelle zu beliebigen Datenbanksystemen mit festem Ausgabeformat; ein Programm, um einen Typen in mehrere zu teilen oder mehrere zu einem zusammenzulegen, um damit interaktiv zu experimentieren; Vergleich von vielen Seriationsergebnissen mit unterschiedlichen Startreihenfolgen, um mittlere Grab- und Typpositionen und deren Streuung zu berechnen; ein Simulationsprogramm, um künstliche Testdaten zur Programmprüfung zu erzeugen.

Die Programme können auf Anfrage kostenlos auf 1600 oder 800 bpi 9-Spur Magnetband in ASCII oder EBCDIC oder auf 8 Zoll Floppy in CP/M oder DEC RX01 Format zugeschickt werden. 5-1/4 Zoll Floppies sind vielleicht zur nächsten Version des Pakets in einer Reihe verbreiteter Mikrocomputerformate und für die MS/DOS und CP/M Betriebssysteme verfügbar.

Pilotinstallationen unterschiedlicher Ausrüstung, die sich nach erfolgreicher Anpassung und Benutzung des Pakets auf der eigenen Anlage als Zweitverteiler für alle weitgehend ähnlichen Systeme zur Verfügung stellen sollen, werden bevorzugt beliefert. Die Originalversion für DEC-Maschinen, die den DECUS Swedish PASCAL Kompiler benutzt, kann in Bonn erhalten werden.

Wie man die Programme erhält:

auf Magnetband:

Schicken Sie uns ein normales 300, oder größeres Magnetband und schreiben Sie, ob Sie ASCII oder EBCDIC und 800 oder 1600 bpi brauchen.

wenn auf Diskette:

Formatieren Sie bitte 5 einseitige einfachdichte 8 Zoll Disketten mit ihrem System. Wir können CP/M, DEC RT-11 und DEC Files-11 Directories erzeugen. Bitte geben sie an, welches erwünscht ist:

Senden Sie das Magnetband/die Diskette an:

Labor für Feldarchäologie
Rheinisches Landesmuseum
Colmantstr. 14
D 5300 Bonn 1

Dies ist das erste aus einer Serie von archäologischen Programmpaketen für die Lösung häufig auftretender Probleme, wobei

effiziente moderne Algorithmen benutzt werden. Alle werden realisiert in gut lesbaren und strukturierten Programmen in PASCAL und C, wobei Portabilität ein bewußtes Ziel ist, daher gedacht für weltweite Verteilung. Zu den Programmen gibt es eine ausführliche Bibliographie in maschinenlesbarer Form, die auch in einer Datenbank abgelegt werden kann. Die Dokumentation und die Benutzerinteraktion ist auf englisch, aber die Interaktionssprache kann bei Installation verändert werden.