Navigating INFO

presented by:
Tim Haithcoat
University of Missouri
Columbia

with materials from:
Environmental Systems Research Institute
Section One

Introduction
Course Goals

- Help you decrease time and resource use in arc/info tasks
- Give you ideas on how to approach complex ARC/INFO applications
- Help you creatively solve problems
- Understand the “inner workings” of database software (like INFO)
- Make info more enjoyable to use - move away from “brute force” solutions, the drudgery of everyday info
- Help you avoid problems resulting from inefficient techniques and database design
- Demonstrate good relational database design
Course Pre-Requisites

- Realization that you spend most of your time in INFO
- That you agree with the following statement:
  - “Once a database is built, 90-100% of your time is spent in INFO.”
- You should have a good grasp of the fundamentals of INFO
  - DISPLAY and LIST data
  - CREATE and MODIFY datafiles
  - Relationship between ARC files and INFO files
  - UPDATE, CALC, and MOVE
  - Basic knowledge of REPROTS, PROGRAMS, and INPUT FORMS (what they are and how they are created)
- Have a desire to improve your understanding of INFO
- A general rule of thumb is one year experience with ARC/INFO
Why Concentrate on INFO?

- Can be horribly slow and frustrating to use
- Once a database is built, typically 90-100% time spent in INFO
- Language of choice for many GIS Algorithms
- Most accessible programming language in ARC/INFO
  - AML cannot access all data and records in a coverage
  - ARC cannot establish powerful relationships like INFO
- Results can usually be obtained in info using crudest methods - not much incentive to experiment
Initially, person B may be only as productive as Person A.

This course is meant to help you reach new “plateaus”.

This course may be too fast for some - don’t despair.

This course may be too slow for some -- later sections are more advanced.

This curve shows how 1 person’s productivity increase over time. This person is not able (or is unwilling) to experiment. Understanding ARC/INFO increases very slowly because of this.

This stairstep shows how experimentation and understanding increase productivity. The stairsteps reflect “plateaus” that are reached as concepts are understood and applied.
Narrative Page

- Instructions that detail the concept found on the concept page
- Narrative text
- Makes the course more ‘stand alone”
- Frees you from taking notes about the details of the concept page
- Take the time to read the narrative page so you know what is there (and what you don’t have to take notes about)
Concept Page

- These pages are used for:
  - Guidance during lecture
  - Key concepts
  - Illustration of specific points
- These concept pages serve as the overheads for the class
Section Two

INFO Key Concepts
With ARC, the use of system resources varies predictably with the number of features.

As number of records increase in INFO, you can count on resource utilization to increase dramatically and vary widely.

Adding more items to an INFO file increase resource use and only adds to the problem. Some ARCH processes are likewise affected by too many items.
Bits are always 0’s and 1’s; therefore, all data is 0’s and 1’s

ASCII Characters (what you see in the system editor) are combinations of 8 0’s and 1’s:
  • 8 bits = 1 byte = 1 ASCII Character
Machine integers are stored as binary (base 2) numbers. (Very compact.)

Floating points are also stored as binary numbers, but are much larger.

Info can perform arithmetic directly on binary and floating point numbers, but must translate ASCII characters (integers and N-Type items).
A 4-byte word is typically used to store real numbers (those with decimal points).

- bits 2-24 = mantissa, 25-32 = exponent
- A 4-byte word translated as floating point is some astronomical number
### Bits & Bytes & Item Definitions

- A 4-byte word translated as machine integer is +/- 2,147,483,648

<table>
<thead>
<tr>
<th>Bytes</th>
<th>INFO Type</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>I</td>
<td>-9 to 99</td>
</tr>
<tr>
<td>2</td>
<td>C</td>
<td>Two characters</td>
</tr>
<tr>
<td>2</td>
<td>B*</td>
<td>+/- 32,767 to +32,768</td>
</tr>
<tr>
<td>4</td>
<td>I</td>
<td>-999 to 9999</td>
</tr>
<tr>
<td>4</td>
<td>C</td>
<td>Four characters</td>
</tr>
<tr>
<td>4</td>
<td>B*</td>
<td>+/- 2,147,483,648</td>
</tr>
<tr>
<td>4</td>
<td>F*</td>
<td>Very large</td>
</tr>
<tr>
<td>4</td>
<td>N</td>
<td>Depends on decimal point,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>range from -.99 to 9999</td>
</tr>
</tbody>
</table>

*machine representation
### Commonly used INFO item types

<table>
<thead>
<tr>
<th>TYPE</th>
<th>STORED AS</th>
<th>WIDTH (BYTES)</th>
<th>LIMITS</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>BINARY</td>
<td>2</td>
<td>+/- 32,767</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4 (max)</td>
<td>+/-2,147,483,648</td>
</tr>
<tr>
<td>I</td>
<td>ASCII</td>
<td>2</td>
<td>-9 to 99</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>-999 to 9999</td>
</tr>
<tr>
<td></td>
<td></td>
<td>16 (max)</td>
<td></td>
</tr>
<tr>
<td>L</td>
<td>ASCII</td>
<td>4</td>
<td>-001 to 9999</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8</td>
<td>-00000001 to 99999999</td>
</tr>
<tr>
<td></td>
<td></td>
<td>16 (max)</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>ASCII</td>
<td>2</td>
<td>-9 to 99</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>-.9 to 9999</td>
</tr>
<tr>
<td></td>
<td></td>
<td>16 (max)</td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>BINARY</td>
<td>4</td>
<td>very small to very large</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8 (max)</td>
<td>unbelievable</td>
</tr>
<tr>
<td>C</td>
<td>ASCII</td>
<td>1 to 4096</td>
<td>any character string</td>
</tr>
<tr>
<td>D</td>
<td>ASCII</td>
<td>8 (only)</td>
<td>Jan 1, 1801 to Dec 31, 9999</td>
</tr>
</tbody>
</table>
### Bits & Bytes & Item Definitions

<table>
<thead>
<tr>
<th>ITEM</th>
<th>As defined</th>
<th>DEFAULT RANGE&lt;sup&gt;1&lt;/sup&gt;</th>
<th>LOWEST RANGE&lt;sup&gt;2&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Definition</td>
<td>Bytes saved</td>
<td>Definition</td>
</tr>
<tr>
<td>MAJOR1</td>
<td>6,6,I</td>
<td>4,6,B</td>
<td>2,5,B</td>
</tr>
<tr>
<td>MINOR1</td>
<td>6,6,I</td>
<td>4,6,B</td>
<td>2,5,B</td>
</tr>
<tr>
<td>MAJOR2</td>
<td>6,6,I</td>
<td>4,6,B</td>
<td>2,5,B</td>
</tr>
<tr>
<td>MINOR2</td>
<td>6,6,I</td>
<td>4,6,B</td>
<td>2,5,B</td>
</tr>
<tr>
<td>TECT-ID</td>
<td>5,5,I</td>
<td>4,5,B</td>
<td>2,5,B</td>
</tr>
<tr>
<td>ALK-ID</td>
<td>5,5,I</td>
<td>4,5,B</td>
<td>2,5,B</td>
</tr>
<tr>
<td>OWNER-NEW</td>
<td>4,4,I</td>
<td>2,4,B</td>
<td>2,4,B</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>12</td>
<td>24</td>
</tr>
</tbody>
</table>

1. If the "as defined" definition was true - numbers between -9999 and 999999
2. If the values were between +/-32,767
ReSelect FOR

vs.

ReSelect BY

ReSelect is an INFO process repeated time and time again.

It’s worth examining its efficiency.
INFO uses QUICKSORT, a general-purpose algorithm

- Avoid sorting ARC files (AAT, PAT, etc.)
- Narrow files sort faster, so minimize the number of items on files and their widths
- Use sorted files to take advantage of reselect by and relate with order; design your database accordingly
- Only selected files (not related files) can be sorted
Each internal variable starts out as 8 bytes
$CHR1, $NUM1, and $DAT1 are all the same buffer space
$CHR1 will overwrite $NUM1 and $DAT, and vice versa
Formatting a variable to be wider than 8 bytes overwrites one or more subsequent variables
The Coverage

Workspace

(INFO)

<table>
<thead>
<tr>
<th>arcdr9</th>
<th>arcnsp</th>
</tr>
</thead>
<tbody>
<tr>
<td>cover.tic</td>
<td>template</td>
</tr>
<tr>
<td>cover.bnd</td>
<td>template</td>
</tr>
</tbody>
</table>

(Coverage)

<table>
<thead>
<tr>
<th>tic</th>
<th>bnd</th>
</tr>
</thead>
</table>

(operating system directory)
(operating system file)
INFO files & SYSTEM files

INFO data file

ITEM1  ITEM2
123    ABCD
222    EFGH
321    IJKLM
9999    NOPQR

System files

ARCnnnNIT

0100100100100
0010010010011
1100100100100
1100100100100

ARCnnnDAT

0100100100100
0010010010011
1100100100100
1100100100100
INFO files & SYSTEM files

INFO program report or form
- SELECT FILE1
- REL FILE2 BY I
- RESLECT FOR
- PROGEAM SECTIO
- IF $1ITEM GT
- CALC ITEM -
- ENDIF

INFO directory file
- FILE1 ARC000
- FILE2 ARC001
- PROG1 ARC002
- PROG2 ARC003
- REP1 ARC004
- FORM1 ARC005

System files
- SELECT FILE1
- REL FILE2 BY I
- RESELECT FOR
- PROGRAM SECTIO
- 0100100100100
- 0010010010011
- 1100100100100
- 1100100100100

System files
- 0100100100100
- 0010010010011
- 1100100100100
- 1100100100100

System files
- 0100100100100
- 0010010010011
- 1100100100100
- 1100100100100

System files
- 0100100100100
- 0010010010011
- 1100100100100
- 1100100100100

System files
- 0100100100100
- 0010010010011
- 1100100100100
- 1100100100100

System files
- 0100100100100
- 0010010010011
- 1100100100100
- 1100100100100

System files
- 0100100100100
- 0010010010011
- 1100100100100
- 1100100100100

System files
- 0100100100100
- 0010010010011
- 1100100100100
- 1100100100100

System files
- 0100100100100
- 0010010010011
- 1100100100100
- 1100100100100

System files
- 0100100100100
- 0010010010011
- 1100100100100
- 1100100100100

System files
- 0100100100100
- 0010010010011
- 1100100100100
- 1100100100100

System files
- 0100100100100
- 0010010010011
- 1100100100100
- 1100100100100

System files
- 0100100100100
- 0010010010011
- 1100100100100
- 1100100100100

System files
- 0100100100100
- 0010010010011
- 1100100100100
- 1100100100100

System files
- 0100100100100
- 0010010010011
- 1100100100100
- 1100100100100

System files
- 0100100100100
- 0010010010011
- 1100100100100
- 1100100100100

System files
- 0100100100100
- 0010010010011
- 1100100100100
- 1100100100100

System files
- 0100100100100
- 0010010010011
- 1100100100100
- 1100100100100
### Internal Files

**INTERNAL FILE:**
DAT File holds the data

- INFO file
- ARCnnnDAT

### External Files

**EXTERNAL FILE:**
DAT File holds a pointer to a system file that holds the data.

- EXT:
  - ARCnnnDAT
  - MYDIR>coverage>PAT

(AARC files are like this.)

- INFO file
- System file
ARC files are externalled files:

- **WORKSPACE**
  - **INFO**
    - TIC
    - DAT
  - **COVERAGE**
    - PAT
    - AAT
    - BND
    - TIC
Program Sections Odd & Even

- Programs entered with the PROGRAM command
- Edited with CHANGWE
- Compiled with COMPILe
- Executed with RUN
- Deleted with ERASE

**PROGRAM SECTION ODD**

- Just like interactive INFO
- Think of it as a “select block”
- Operates on whole file at once
  - PROGRAM T$CALC
  - SEL LINCOVER.AAT
  - CALC ITEM1 = ITEM1 /10
  - CALC ITEM2 = ITEM1 + 100

**If LINEV COVER.AAT** had 1000 records, each CALC statement executes 1000 times.
  - In this example, we do 2000 access of the file (2 * 1000)
Program Sections Odd & Even

**PROGRAM SECTION EVEN**

- Cannot occur during an interactive session
- think of it as a big “do loop:
- Operates on one record at a time
  1. PROGRAM T$CALC
  2. SEL LINECOVER.AAT
  3. PROGRAM SECTION EVEN
  4. CALC ITEM1 = ITEM1 /10
  5. CALC ITEM2 = ITEM1 + 10
  6. <carriage return>
  7. COMPILE T$CALC
  8. RUN T$CALC
- Lines 4 and 5 are in a big “do loop”
- At the start of program execution, the first selected record is passed through the even section
- The overall effect is that we only do 1000 accesses of the file
- Obviously, we have just saved a bunch of time