The Representation and Storage of Combinatorial Block Designs

Hatem Nassrat

April 15, 2009
Combinatorial Block Designs

Project Intro

External Representation

Design Database

Web Interface

System Deployment

System Overview

Conclusions
What are Block Designs?

- Multiset of subsets of points

\[ D = (B, V) ; B = \{ b \mid b \subset V \} \]
What are Block Designs?

- Multiset of subsets of points

\[ D = (B, V) ; \quad B = \{ b \mid b \subset V \} \]

- Example (Fano Plane)

\[ V = [0, 1, 2, 3, 4, 5, 6] \]
\[ B = [ [0, 1, 2], [0, 3, 4], [0, 5, 6], [1, 3, 5], [1, 4, 6], [2, 3, 6], [2, 4, 5] ] \]
What are Block Designs?

- Multiset of subsets of points
  \[ D = (B, V); \ B = \{ b | b \subset V \} \]

- Example (Fano Plane)
  \[ V = [0, 1, 2, 3, 4, 5, 6] \]
  \[ B = [ [0, 1, 2], [0, 3, 4], [0, 5, 6], [1, 3, 5], [1, 4, 6], [2, 3, 6], [2, 4, 5] ] \]

- Useful in experimental design, finite geometry, software testing, cryptography, error correcting codes.
What are Block Designs?

► Multiset of subsets of points

\[ D = (B, V) ; B = \{ b | b \subset V \} \]

► Example (Fano Plane)

\[ V = [0, 1, 2, 3, 4, 5, 6] \]
\[ B = [ [0, 1, 2], [0, 3, 4], [0, 5, 6], [1, 3, 5], [1, 4, 6], [2, 3, 6], [2, 4, 5] ] \]

► Useful in experimental design, finite geometry, software testing, cryptography, error correcting codes.

► Hamming codes were discovered as a block design 5 years earlier [Cam08].
Project Phases

- External Representation of Block Designs
- Block Design Database
- Web Interface
Project Motivation

- Searching for block designs
- Central block design repository
- Concise readable representation
What is the Ext Rep?

- Method of communicating designs
- Communicating design properties
- Invented at the Queen Mary University of London
- Platform Independent
- User Agents = Human | Software
- Ext Rep v2 published mid 2004
Ext Rep v3

- v2 in XML
- v3 in JSON
- Enhanced Structure
- Enhanced Functionality

```
<blocks ordered="true">
  <block>0</block> <block>1</block> <block>2</block>
  <block>0</block> <block>3</block> <block>4</block>
  <block>0</block> <block>5</block> <block>6</block>
  <block>1</block> <block>3</block> <block>5</block>
  <block>1</block> <block>4</block> <block>6</block>
  <block>2</block> <block>3</block> <block>6</block>
</blocks>
```

(a) v2 [CDMS03]

```
"blocks" : [
  [0, 1, 2], [0, 3, 4], [0, 5, 6],
  [1, 3, 5], [1, 4, 6], [2, 3, 6],
  [2, 4, 5]
]
```

(b) v3 [DN08]

Figure: blocks in ExtRep versions
Implementing the Conversion

- Automated Conversion
- Custom JSON Parser
  - Stream Parsing
  - Flexible internal representation
  - YAJL, callback based C library
DBMS Options
- Hierarchical DBMS HDF5
- Relational DBMS PostgreSQL

PostgreSQL Prevailed
- Faster reads
- Simpler to Implement
Designing the Database

- Ext Rep structure is Tree Like
- Table per optional Sub-Tree
- Foreign Keys
- List → Many-to-Many
- Automating the Transformation
Populating the Database

- Callback Parsing
- Isomorphic Rejection
- Categorizing Designs
Query Engine

- Encapsulated Design
- Query Language
- User/Session Management

Figure: Query Example
Web Interface

- Pylons Framework
- Query Interface
- Interface with Query Engine
- Displaying Designs on the Web
Web Interface

- Pylons Framework
- Query Interface
- Interface with Query Engine
- Displaying Designs on the Web
Web Interface

- Pylons Framework
- Query Interface
- Interface with Query Engine
- Displaying Designs on the Web
Web Interface

- Pylons Framework
- Query Interface
- Interface with Query Engine
- Displaying Designs on the Web
Web Interface

- Pylons Framework
- Query Interface
- Interface with Query Engine
- Displaying Designs on the Web
Web Interface

- Pylons Framework
- Query Interface
- Interface with Query Engine
- Displaying Designs on the Web
System Deployment

- Designed for Unix
- Apache2 w/ Mod Wsgi
- Load testing
System Overview & Usage

Figure: Design DB Overview
## System Overview & Usage

**Figure:** Web Visit Aggregates for Weekdays in January 2009

<table>
<thead>
<tr>
<th>Day</th>
<th>Pages</th>
<th>Hits</th>
<th>Bandwidth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mon</td>
<td>590.75</td>
<td>715.25</td>
<td>9.10 MB</td>
</tr>
<tr>
<td>Tue</td>
<td>570.25</td>
<td>688.50</td>
<td>12.54 MB</td>
</tr>
<tr>
<td>Wed</td>
<td>602</td>
<td>721.75</td>
<td>10.97 MB</td>
</tr>
<tr>
<td>Thu</td>
<td>577.20</td>
<td>694.20</td>
<td>8.77 MB</td>
</tr>
<tr>
<td>Fri</td>
<td>565.80</td>
<td>648</td>
<td>7.61 MB</td>
</tr>
<tr>
<td>Sat</td>
<td>534.25</td>
<td>564</td>
<td>3.31 MB</td>
</tr>
<tr>
<td>Sun</td>
<td>536.75</td>
<td>572.75</td>
<td>3.89 MB</td>
</tr>
</tbody>
</table>
## System Overview & Usage

**Figure:** Web Visit Durations for January 2009

<table>
<thead>
<tr>
<th>Visits duration</th>
<th>Number of visits</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>0s-30s</td>
<td>236</td>
<td>57 %</td>
</tr>
<tr>
<td>30s-2mn</td>
<td>40</td>
<td>9.6 %</td>
</tr>
<tr>
<td>2mn-5mn</td>
<td>23</td>
<td>5.5 %</td>
</tr>
<tr>
<td>5mn-15mn</td>
<td>16</td>
<td>3.8 %</td>
</tr>
<tr>
<td>15mn-30mn</td>
<td>4</td>
<td>0.9 %</td>
</tr>
<tr>
<td>30mn-1h</td>
<td>8</td>
<td>1.9 %</td>
</tr>
<tr>
<td>1h+</td>
<td>84</td>
<td>20.2 %</td>
</tr>
<tr>
<td>Unknown</td>
<td>3</td>
<td>0.7 %</td>
</tr>
<tr>
<td>Countries</td>
<td>Pages</td>
<td>Hits</td>
</tr>
<tr>
<td>-------------------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td>Unknown</td>
<td>ip</td>
<td>10068</td>
</tr>
<tr>
<td>United States</td>
<td>us</td>
<td>6313</td>
</tr>
<tr>
<td>Great Britain</td>
<td>gb</td>
<td>155</td>
</tr>
<tr>
<td>European country</td>
<td>eu</td>
<td>128</td>
</tr>
<tr>
<td>Canada</td>
<td>ca</td>
<td>107</td>
</tr>
<tr>
<td>Italy</td>
<td>it</td>
<td>75</td>
</tr>
<tr>
<td>Germany</td>
<td>de</td>
<td>53</td>
</tr>
<tr>
<td>Australia</td>
<td>au</td>
<td>27</td>
</tr>
<tr>
<td>Spain</td>
<td>es</td>
<td>18</td>
</tr>
<tr>
<td>France</td>
<td>fr</td>
<td>15</td>
</tr>
<tr>
<td>Others</td>
<td></td>
<td>72</td>
</tr>
</tbody>
</table>

**Figure**: User Countries - January 2009
Conclusion

- Combinatorial & Statistical Designs
  - Ext Rep v3
  - Design DB
  - Web Interface
- Computer Science
  - System Integration
  - Source Code
  - Evaluation of Open Source Software
Questions?


Future Plans

- Ext Rep Extension
- RDBMS Alternatives
- Interface Personalization
- High Performance
- System Upgrade (Python, Pylons, SqlAlchemy, YAJL)
- Application Programmers Interface
- Accepting Contributions (Uploading)