Querying and Visualizing Electronic Health Records

Taowei David Wang
Catherine Plaisant
Ben Shneiderman
Stanley Lam

Human-Computer Interaction Lab, University of Maryland
LifeLines: Overview of Patient Record

Linda Simpson
Female 40

LifeLine
Notes
Tobacco
Depression
Lyme
Arthritis

Obesity
Checkup
Checkup
Checkup
Checkup

Arrhythmia
Pneumonia
Knee Pain
Diabetes
Pregnancy

Hosp.
Appendectomy

Tests
Blood EKG
ENo
Knee Surgery

Med.
Prozac

Blood
Blood
Blood
Blood

Other
Physical Therapy

Immune
TB Test
Tetanus

CHI96 AMIA98 - www.cs.umd.edu/hcil/lifelines
Single record $\rightarrow$ Millions of records

- Large databases of Electronic Health Records (EHRs):
  - Observational studies
  - Recruitment for clinical trials
  - Hospital metrics
  - Alarm design and testing
  - etc.

Often involves temporal comparison relative to an important event (e.g. heart attack, start of a treatment, 1st diagnosis of cancer)
Lifelines2
Sets of patient records
Disclaimer: de-identified data
Lifelines2

- Introduce powerful combination of simple operations: Align, Rank, Filter, and Summarize
  - Multiple records simultaneously visible
  - Align by sentinel events
  - Rank by frequency
  - Filter by events
  - Summaries

- Focus on Point Events
  - Diagnosis, treatments, etc.

- Measure Benefits of Alignment
Context

Search millions of records

Interactive visualization of results

LifeLines2
Context

Search
millions of records

Interactive
visualization of results

Writing SQL not an option!

LifeLines2
Writing SQL not an option!

```
SELECT TOP 200 *
FROM
    contrast contrast_1,
    contrast contrast_2,
    contrast contrast_3
WHERE
    1=1 AND
    [contrast_1].[Lab Test] = 'CREATE' AND
    [contrast_2].[Value] BETWEEN '0.6' AND '1.2' AND
    [contrast_2].[Lab Accessioning Time] < [contrast_1].[Radiology Exam Time] AND
    [contrast_2].[Lab Accessioning Time] > dateadd(Day, -2, [contrast_1].[Radiology Exam Time]) AND
    [contrast_1].[Account] = [contrast_2].[Account] AND
    [contrast_3].[Value] > ([contrast_2].[Value] * (1 + 50 * .01)) AND
    [contrast_3].[Value] > ([contrast_2].[Value] + 1) AND
    [contrast_3].[Lab Accessioning Time] > [contrast_1].[Radiology Exam Time] AND
    [contrast_3].[Lab Accessioning Time] < dateadd(Day, 5, [contrast_1].[Radiology Exam Time]) AND
    [contrast_2].[Account] = [contrast_3].[Account]
```
• Sample of Related work
• Demo
• (Quick) Report on studies
• Ongoing & Future work
Static views

Powsner & Tufte, 1994

Lexis diagrams (Bertin)
Lifelines and improvements
Overview of categorical and/or numerical data (semantic zoom)
Lifelines and improvements
Overview of categorical and/or numerical data (semantic zoom)

Plaisant, CHI 96, AMIA 98

i2b2 (Murphy, AMIA 07)
Lifelines and improvements
Overview of categorical and/or numerical data (semantic zoom)

Plaisant CHI 96, AMIA 98
Bade, CHI 2004

i2b2 (Murphy, AMIA)
Temporal Visualizations

Spiral Graph: Weber, 01 (based on Carlis, UIST 89)

ExperiScope (Guimbretiere, CHI 07)

Periodic data

One of many example of manual alignment
PatternFinder
Specification of complex temporal queries on categorical data

Ball and chain display of matches
• Related work 4
• **Demo LifeLines2**
• Report on studies
• Ongoing & Future work
Scenario: Study relationship between asthma and pneumonia
1) Run query
Scenario: Study relationship between asthma and pneumonia

1) Run query Find all patients who have both Asthma and Pneumonia diagnoses
Scenario: Study relationship between asthma and pneumonia

1) Run query Find all patients who have both Asthma and Pneumonia diagnoses

2) Review results
Demo
User Studies
Two user studies

• Controlled experiment
  (some training, measure speed and error)

• Domain expert qualitative study
  (no training, think aloud, discussion)
Two user studies

• Controlled experiment (some training, measure speed and error)
  • Benefit of alignment: **YES** (Significant improvement on complex tasks)
  • 20 participants: grad students
    Data: synthetic student record data
    Tasks checked as domain independent

• Domain expert qualitative study (no training, think aloud, discussion)
Two user studies

• Controlled experiment (some training, measure speed and error)
  • Benefit of alignment: **YES** (Significant improvement on complex tasks)

  • 20 participants: grad students
    Data: synthetic student record data
    Tasks checked as domain independent

• Domain expert qualitative study (no training, think aloud, discussion)

  • Learnability **GOOD** (challenges with interpretation of data more than UI)
  • General feedback and suggestions

  • 4 participants: nurse, physician, 2 prof. of nursing
  • All experienced with EHR and medical research

---

![Task Completion Time Chart](chart.png)

**Task Completion Time**

- **With Alignment**
- **Without Alignment**

<table>
<thead>
<tr>
<th>Tasks</th>
<th>1</th>
<th>2*</th>
<th>3</th>
<th>4*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time (seconds)</td>
<td>40</td>
<td>60</td>
<td>40</td>
<td>60</td>
</tr>
</tbody>
</table>

* Asterisk denotes tasks with significant improvement with alignment.
Context

Search millions of records

Interactive visualization of results

LifeLines2
Context

Search millions of records

Interactive visualization of results

PatternFinder in Amalga (formerly Azyxxi)

LifeLines2
Combine Alignment with PatternFinder

Results for: Contrast and Creatinine AMIA
Aligned by: Sentinel Event

-2 day  -1 day  0 day  +1 day  +2 day  +3 day

Paper submitted

Washington Hospital Center
Integrate Align-Rank-Filter in i2b2

Harvard Medical School, Partners HealthCare
In summary...

- **Align Rank, Filter, and Summarize**
  Powerful combination of simple operations to explore temporal categorical data (events)

- **Performance benefit of alignment:** significant

- **Impact:** Integration in 2 large operational EHR systems

- **Many applicable domains:**
  - Highway incident log
  - Student records
  - Web logs
  - Vehicle fleet records
Thank you!

www.cs.umd.edu/hcil/lifelines2
Fading intervals
For all patients who had a radiology exam using contrast, give me all their lab results which included the string “creat”.

Data exported from Azyxxi (and deidentified)
<table>
<thead>
<tr>
<th>Records</th>
<th>3398/3398</th>
</tr>
</thead>
</table>

| 0000537 | CREATE-H Radiology Contract |
| 0000538 | CREATE-H Radiology Contract |
| 0000539 | CREATE-H Radiology Contract |
| 0000540 | CREATE-H Radiology Contract |
| 0000541 | CREATE-H Radiology Contract |
| 0000542 | CREATE-H Radiology Contract |
| 0000543 | CREATE-H Radiology Contract |
| 0000544 | CREATE-H Radiology Contract |
| 0000545 | CREATE-H Radiology Contract |
| 0000546 | CREATE-H Radiology Contract |
| 0000547 | CREATE-H Radiology Contract |
| 0000548 | CREATE-H Radiology Contract |
| 0000549 | CREATE-H Radiology Contract |
| 0000550 | CREATE-H Radiology Contract |
| 0000551 | CREATE-H Radiology Contract |
| 0000552 | CREATE-H Radiology Contract |
| 0000553 | CREATE-H Radiology Contract |
| 0000554 | CREATE-H Radiology Contract |
| 0000555 | CREATE-H Radiology Contract |

**Legend**:
- CREATE-H
- CREATE-L
- CREATE-R
- CREATE-T
Qualitative Evaluation Results

• Visual representation and ARF easily understood
  – Readily apply without training
• 2/4 participants saw immediate utility in their own medical research

• IV alters perception, interpretation of data
• IV can be confusing
  – Did not trust how intervals are assigned
  – Interpreted the interval literally
• Feedback on improving ARF
Quant. Evaluation Sample Data

Back to Evaluations
(Alignment vs. No Alignment)

**Task Completion Time**

<table>
<thead>
<tr>
<th>Tasks</th>
<th>With Alignment</th>
<th>Without Alignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20</td>
<td>40</td>
</tr>
<tr>
<td>2</td>
<td>60</td>
<td>120</td>
</tr>
<tr>
<td>3</td>
<td>80</td>
<td>100</td>
</tr>
<tr>
<td>4</td>
<td>100</td>
<td>120</td>
</tr>
</tbody>
</table>

**Task Error Rate**

<table>
<thead>
<tr>
<th>Rate</th>
<th>With Alignment</th>
<th>Without Alignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.1</td>
<td>0.2</td>
</tr>
<tr>
<td>0.1</td>
<td>0.3</td>
<td>0.4</td>
</tr>
<tr>
<td>0.2</td>
<td>0.5</td>
<td>0.6</td>
</tr>
<tr>
<td>0.3</td>
<td>0.7</td>
<td></td>
</tr>
</tbody>
</table>

**Task Error Size**

<table>
<thead>
<tr>
<th>Average Error Size</th>
<th>With Alignment</th>
<th>Without Alignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.1</td>
<td>0.2</td>
</tr>
<tr>
<td>0.1</td>
<td>0.3</td>
<td>0.4</td>
</tr>
<tr>
<td>0.2</td>
<td>0.5</td>
<td>0.6</td>
</tr>
<tr>
<td>0.3</td>
<td>0.7</td>
<td></td>
</tr>
</tbody>
</table>
(IV vs. no IV)

**Task Completion Time**

- **With IV**
- **Without IV**

**Task Error Rate**

- **With IV**
- **Without IV**

**Task Error Size**

- **With IV**
- **Without IV**

**Average Error Size**

- **With IV**
- **Without IV**
Quantitative Evaluation Tasks

· **Task 1**: How many students submitted a paper within 1 month after proposal? (5 records)

· **Task 2**: How many students submitted a paper within 1 month after proposal? (20 records)

· **Task 3**: How many students published at least 3 papers between proposal and defense?

· **Task 4**: What occurred most often within a month of a student’s 1st paper submission?

For interval of validity

· **Task 5**: Assuming a class lasts 3 months, how many students proposed while they were taking a class?

· **Task 6**: Assuming a class lasts 3 months, and it takes 2 months to prepare for proposal, how many students were preparing for proposal while taking a class?
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Severe headache</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seizure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smoker</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thoracic outlet syndrome</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depression</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Allergies       |       |       |       |       |       | asa - rash | codeine - angioedema | penicillin - anaphylaxis | toradol - angioedema |

| Diagnosis        |       |       |       |       |       | Migraine | Depressive disorder | Smoker | Seizure disorder | Possible irritable bowel syndrome |

| Complaints       |       | Headache | Migraine | Headache | Migraine | Left Lower Quadrant pain | Constipation |

| Lab-Path         | Dilantin(Phenytoin) | Liver profile |

| Imaging          |       | Athenolol | Propanolol | Prochlorperazine | Promethazine HCL | Butalbital/apap/caffeine tab | Chloral hydrate | Dilantin | Paxil | Chloral hydrate | Dilantin | Meperidine | Meperidine | Nortriptyline HCl |

| Medications      |       |           |           |           |           |           |               |            |       |               |            |             |             |                 |
Medications examples

- Drug A: Order, Register, Dispensed
- Drug B: Order, Register, Dispensed
- Stop request, Stop

Today:
- Athenolol: (Scheduled stop date)
- Chlorol hydrate: Disp, Disp
- Dilantin: Disp, Disp, Disp
- Not requesting refill
- Aleve: (Thickness for dosage)
Summarization

- Important issue because of large time range
- Recursive aggregation mechanism
  (set of events) --- (summary event)
### Records 37/37 Rank by [# of Pneumonia and Influences]

<table>
<thead>
<tr>
<th>Record</th>
<th>Year</th>
<th>Asthma</th>
<th>Pneumonia and infl</th>
<th>Viral pneumonia</th>
</tr>
</thead>
<tbody>
<tr>
<td>000000001</td>
<td>1984</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000112</td>
<td>1985</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000112</td>
<td>1986</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000112</td>
<td>1987</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000112</td>
<td>1988</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000112</td>
<td>1989</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000112</td>
<td>1990</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000112</td>
<td>1991</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000112</td>
<td>1992</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000112</td>
<td>1993</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000112</td>
<td>1994</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000112</td>
<td>1995</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000112</td>
<td>1996</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000112</td>
<td>1997</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000112</td>
<td>1998</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000112</td>
<td>1999</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000112</td>
<td>2000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000112</td>
<td>2001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000112</td>
<td>2002</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000112</td>
<td>2003</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000112</td>
<td>2004</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000112</td>
<td>2005</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000112</td>
<td>2006</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000112</td>
<td>2007</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000112</td>
<td>2008</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000112</td>
<td>2009</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000112</td>
<td>2010</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000112</td>
<td>2011</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000112</td>
<td>2012</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000112</td>
<td>2013</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000112</td>
<td>2014</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000112</td>
<td>2015</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000112</td>
<td>2016</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000112</td>
<td>2017</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000112</td>
<td>2018</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000112</td>
<td>2019</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000112</td>
<td>2020</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000112</td>
<td>2021</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Ranked by [# of Pneumonia and Influences]**
Also fewer, smaller errors
With alignment
Not really useful if data fit in one screen
or if still learning i.e. best strategy not immediate
TimeSearcher
Dynamic queries on numerical temporal data

Hochheiser Infovis04

Buono VDA05
Specification of temporal abstractions
To reason/query with them

Shahar 1999

No focus on interaction

Post 2007
Combine alignment with PatternFinder

Collaboration with Washington Hospital Center
(original developers of Azyxxi, now Microsoft Amalga)

• Connected to real time database
• Search UI was designed to match style of existing search interface

• Alignment integrated
• Put on desk of first users soon
Measured benefits over tabular display

www.cs.umd.edu/hcil/lifelines
PatternFinder
Specification of complex temporal queries on

Patients with increasing dosages of Remeron followed by a heart attack within 180 days

[Fails et al. VAST06]
Spiral Graph: Weber 01 (based on Carlis UIST 89)

Periodic data

ThemeRiver (Havre, Infovis00)

Experiscope (Guimbretiere, CHI07)

One of many example of manual alignment