Being Provocative in Preference Handling or Why OS Should be the Emblem of our Workshops

Carmel Domshlak (Technion, Israel)



"Case Study"

Some preference information

Evaluating alternatives in terms of user preferences.

Database







List of alternatives that suit user's preference best.

 $\mathbf{X} = \{X_1, \dots, X_n\}$ $\mathcal{X} = \times Dom(X_i) \equiv \Omega$

Characteristic Properties

- I. Knowing the ordinal preferences suffices
- 2. Multi-attribute description of the alternatives

Direct assessment of a preference ranking is typically infeasible as the size of the tuple space is exponential in n

- 3. Lay users and (possibly) NO decision analyst around
- 4. On-line decisions

Statements of ordinal preference

Each query consists of a set of (qualitative? quantifying? hybrid?) statements communicating us some information about the ordinal preferences of the user.

I prefer Continental to Delta ...

Soup is more important to me than desert ...

For me, the value of automatic transmission in a mini-van is \$1000...





General Framework



Good news

- From philosophical logic to computer science
- Numerous "toolkits" that
 - support certain forms of preference statements
 - suggest concrete semantics for their interpretation
 - provide some computational means for reasoning about the interpretations

Example: Graphical Models



- I. Adopt a concrete semantics for statement interpretation.
- 2. Identify useful notions of purely qualitative, possibly conditional preferential independence.
- 3. Use preferential independence to define graphical models for preference representation.
- 4. Exploit the graphical core of the models to achieve computational efficiency.

So the good news are ...

- From philosophical logic to computer science
- Numerous "toolkits" that
 - support certain forms of preference statements
 - suggest concrete semantics for their interpretation, and
 - provide some computational means for reasoning about the interpretations

Paradoxical Situation

- The topic of decision/choice support is a priori of high applicative potential
- Various models have been developed to provide semantic and computational infrastructure
- No applications that are based on non-trivial preference specification
- Not even attempts ...

- The users are too lazy to provide non-trivial preference information
 - Convincing?

- The users are too lazy to provide non-trivial preference information
 - Convincing?

Reminds me some story on Alexander G. Bell and American Telegraph ...



- The users are too lazy to provide non-trivial preference information
 - Convincing?

Reminds me some story on Alexander G. Bell and American Telegraph ...



- 8-10 years ago I had hard times to return a battle ...
- ... but during the last decade, the argument has lost its power!
 - Verbal user opinions on virtually everything (and for free!)
 - Systems for emotion detection in texts

- The users are too lazy to provide non-trivial preference information
 - Convincing?
- The preferences of the users are typically simple
 - Convincing?

Empirical evaluation of ordinal preferences (with T. Joachims)

- EachMovie data set
 - Ratings of 1628 movies by 72916 users.
 - Six point scale: 0 ("worst") 5 ("best")
 - Movie attributes: Decade and 10 genre categories.
- (So far inherent) problem: No generalizing statements.
- Solution: "Reveal" these statements from the data.
 - Learn rules using C4.5 decision tree learning.
 - Pick probabilistically most significant rules.

"Preference statements"

```
B_decade = 90s
B_Art_Foreign = 1
B_Family = 0
B_Romance = 0
-> user prefers movie A over movie B [100%]
```

The user doesn't like foreign films from the 90s that are not romance or family movies.

```
A_decade = 80s
A_Thriller = 1
B_Classic = 0
B_Horror = 1
-> user prefers movie A over movie B [96.2%]
```

The user prefers thrillers from the 80s to non-classic horror movies.

Some test results



- The users are too lazy to provide non-trivial preference information
 - Convincing?
- The preferences of the users are typically simple
 - Convincing?
- The models we suggest do not suit the real needs
 - May be ... but what is the evidence for that?



- From only philosophical logic to computer science
- Numerous "toolkits" that
 - support certain forms of preference statements
 - suggest concrete semantics for their interpretation
 - provide some computational means for reasoning about the interpretations

Where do we stand ...

Given a "toolkit" for reasoning about human preference statements, it is hard to say *a priori*

- to which application domains (if any) the toolkit will apply well?
- to what degree the toolkit covers the needs of the domain?
- are there better toolkits for this domain?

Where do we stand ...

Given a "toolkit" for reasoning about human preference statements, it is hard to say *a priori*

- to which application domains (if any) the toolkit will apply well?
- to what degree the toolkit covers the needs of the domain?
- are there better toolkits for this domain?



No data \mapsto No evaluation \mapsto No data ...

What I think we need (urgently)

Developing mathematical models for *Information Retrieval* without having some benchmark information sets and queries would probably not get too far

We are working on a similar (but seemingly more complicated) problem!

We need:

- benchmarks
- comparative evaluations
- "competitions"

Going empirical! Area success = F(degree of being empirical)

Examples from Artificial Intelligence

- + Automated planning
- Non-monotonic reasoning
- ± Probabilistic reasoning
- **±** Machine learning

Going empirical! Area success = F(degree of being empirical)

Examples from Artificial Intelligence

- + Automated planning
- Non-monotonic reasoning
- **±** Probabilistic reasoning
- **±** Machine learning

Examples from Information Processing

Information retrieval

+ in tasks supported by benchmarks, - everywhere else

Natural language processing

- NL understanding, but + in parsing, and even (recently) text entailment

Databases

+ in SQL queries processing, - (so far) in XML query processing

The challenge of benchmark engineering We try modeling a cognitive paradigm!

- (Unfortunately) our problems are closer to IR and NLP, rather than to Automated Planning or CSP
- Major problem: Absence of real-world data
 - What we did with Thorsten Joachims was a "reasonable and justifiable fake".

The challenge of benchmark engineering We try modeling a cognitive paradigm!

- (Unfortunately) our problems are closer to IR and NLP, rather than to Automated Planning or CSP
- Major problem: Absence of real-world data
 - What we did with Thorsten Joachims was a "reasonable and justifiable fake".
- Engineering "synthetic benchmarks"
 - What kind of data we want? How synthetic can it really be?
 - Requires cooperation with experimental psychology
 - Requires funding!

The challenge of benchmark engineering

rather t

Major p

Most benchmarks that pushed forward human-related computation areas came from (sometimes massive) funding from major funding agencies - DARPA ↔ Classical IR

- Enginee
 ARDA ↔ NLP / Question answering
 EU ↔ Text entailment

- What kind of data we want?

- Requires funding!

The news we would like to hear



- The user should be able to provide arbitrary preference information she finds *natural*
- Reasoning about preference information should be completely non-parametric, that is free of the explicit assumptions about the structure of the user preferences.
- Reasoning about preference information should be computationally efficient, even in cases where user preferences pose no significant independence structure on the attributes in use

This is not a hopeless goal!

- Initial but very promising results in [DT05-07]
- Promise of empirical proofs of suitability of certain toolkits to specific problem domains
- Promise of combining various toolkits based on the (automatically identified!) nature of the statements
- Bringing in understandings from social sciences
- Adapting mathematical and computational techniques from everywhere
 - But only if experimentally or empirically useful!

The news we would like to hear



- The user should be able to provide arbitrary preference information she finds *natural*
- Reasoning about preference information should be completely non-parametric, that is free of the explicit assumptions about the structure of the user preferences.
- Reasoning about preference information should be computationally efficient, even in cases where user preferences pose no significant independence structure on the attributes in use

This is not a hopeless goal, but we should start working towards it!

Let us not be Brezhnev, but his aide!



"Brezhnev begins his official speech opening the 1980 Olympics: 'O! O! O!' His aide interrupts him with a whisper: 'The speech starts below, Leonid Ilich. That is the Olympic symbol.'"

Photos from the Workshop on Handling Preferences, 2015

