Intelligent Support for E-Commerce

Pádraig Cunningham
Department of Computer Science
Trinity College Dublin
Electronic Commerce Applications

- Existing processes re-implemented online
- New models of selling enabled by *electronic* context

*What is special about ‘electronic’ commerce?*
Outline

• Overview of E-commerce
• Existing processes implemented online
  – CBR provides intelligent sales support, fulfils role of sales assistant
  – Product Recommendation
• New possibilities created by electronic context
  – What’s new with e-commerce?
• Use the data - Lazy Learning
• Conclusions
Overview of e-commerce

• Currently
  – 80% business to business
  – 20% business to customer

• “Disintermediation”
  – selling direct to customer will remove some intermediaries
  – new types of intermediaries will emerge - brokers
    • personalogic.com (product recommendation / selection)
    • acses.com (bargain finder)
    • *these can/should be intelligent systems*
Product Types

- Tangible products get media attention
  - computers, books, food/drink, flowers, cars
- Intangible products have probably better potential
  - can be delivered over the net
    - Software, Entertainment, Gambling, Travel services, Financial services, Health services,
    - Pornography probably 3rd behind computers and travel.

Based on OECD Report 1999
Intelligent Systems in e-commerce

• Two main areas:
  – Negotiation
    • Sales Agents
    • Personal Assistants
  – Recommending
    • Determining the correct product for the customer
    • Anonymous or Identified Interaction?
The Buying Process

1. Identification of Needs

2. Buying Decision
   2.1. Information Search
   2.2. Information Evaluation
   2.3. Dealing

3. Purchase

4. Product Service and Evaluation
   (Wilke, 1998)
Related electronic sales process

**Websell Model**
1. Express Demands: Translate user requirements into a query
2. Retrieve appropriate product offers
3. Adapt to meet requirements

(Wilke, 1998)
Successful CBR systems online

• Virtual Property Letting Agent
  – http://www.hookemacdonald.ie/

• Last Minute Travel Bargains
  – http://www.reiseboerse.com/

User is anonymous
Last Minute Travel

Angaben zur gesuchten Last Minute Reise

<table>
<thead>
<tr>
<th>Abflug - Ort</th>
<th>Frankfurt</th>
<th>Reise - Ziel</th>
<th>Mallorca</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abflug - Datum</td>
<td>30.05.1999</td>
<td>Reisedauer in Tagen</td>
<td>7</td>
</tr>
<tr>
<td>Zimmer - Art</td>
<td>Doppelzimmer</td>
<td>Verpflegung</td>
<td>Halbpension</td>
</tr>
<tr>
<td>Angebots - Art</td>
<td>Flug und Hotel</td>
<td>Hier finden Sie Reisen mit Beginn in den nächsten 14 Tagen. Für spätere Termine bitte die Reise Boerse benutzen. Joker steht für &quot;egal wohin - nur weg&quot;.</td>
<td></td>
</tr>
</tbody>
</table>

Reise suchen
Classification from Patti Maes:-

- Feature-based filtering
  - this describes ‘retrieval-only’ CBR systems
  - $k$-Nearest Neighbour matching
- Automatic Collaborative filtering (ACF)
- Constraint-based filtering (CSP)
  - feature-based
  - more sophisticated recommendation than using similarity?
Selecting a laptop

ICCBR August 1999
Asks 18 questions to select between 604 computers.

<table>
<thead>
<tr>
<th>Number of Computers</th>
<th>Total</th>
<th>604</th>
</tr>
</thead>
<tbody>
<tr>
<td>In Your List</td>
<td></td>
<td>32</td>
</tr>
</tbody>
</table>

**RETURN TO START**

1. General Info
2. Technical
3. Multimedia
4. Peripherals
5. Overall Opinion

**YOUR RESULTS**

1. **Compaq Presario 150 with 380 MHz K6-2, 4GB HD, 64MB RAM**

   - Score: 100%
   - Lowest Price Found: $1,899
   - Weight: 5.0 lbs
   - Modern Included: Yes
   - CD-ROM Drive Included: No Data

   Compare With Another Computer
Electronic sales process

1. Express demands
2. Retrieve products
3. Customise products
4. Present products
5. Order processing

Simple electronic sales process

1. Express demands
2. Customer DB
3. Evaluate products
4. Order form

(Wilke, 1998)
What's the difference between $k$-NN and CSP based recommendation?
What’s the difference between $k$-NN and CSP based recommendation?

- Both are feature-based
- CSP
  - offers more sophisticated recommending?
  - harder to set up?
  - asks too many questions?
    - 18 questions to select between 604 laptops
- There are better ways (from ML) to do this
  - Incremental-CBR
    - asks *discriminating* questions
      - (Cunningham, Smyth & Bonzano 1998; Smyth & Cunningham, 1995)
Selling Configurable Products

- **Product Complexity**
  - Standard Products
    - e.g. CDs, books
  - Pre-configured Products
    - e.g. package holidays, standard computers
  - Configurable Products
    - e.g. financial products, technical systems

Recommendation techniques adequate
Requires customization
e.g. selling HVAC systems

- Western Air, HVAC system
  - Watson & Gardingen 1998, 1999
  - supporting sale of heating, ventilation and air-conditioning systems
  - CB of 10,000 records
  - distributed retrieval
  - manual adaptation

- More than just a Recommender system
The story so far

- Recommender systems
  - case-based
  - constraint based
- HVAC sales support system
  - on-line support for configuration of complex products

*existing techniques used online*
What’s special about e-commerce?

• Well, it’s the ‘e’ part obviously!
  – Communication
  – Data

• What kinds of new applications are possible?
  – Systems discussed so far do not exploit available data on user.
Aside: On Internet Advertising

- Traditionally in advertising:
  - 50% of the advertising budget is wasted
  - the problem is knowing what 50% ?

- Internet advertising:
  - advertisers pay for response, mouse clicks
    - doubleclick.com
  - closing the loop only possible in an online context
  - a web-empowered approach
Automatic Collaborative Filtering (ACF)

Machine Learning (McGraw-Hill Series in Computer Science)
by Tom M. Mitchell, Thomas M. Mitchell
Our Price: $85.15
Availability: Usually ships within 24 hours.

Customers who bought this book also bought:

• Reinforcement Learning: An Introduction; R. S. Sutton, A. G. Barto
• Advances in Knowledge Discovery and Data Mining; U. M. Fayyad, et al
• Probabilistic Reasoning in Intelligent Systems; J. Pearl
ACF: How it works

• Users 1, 2 and 3 in same group
  – \{A,B,C\} in common
• D & E can be recommended to User 1 based on this shared interest
• Recommendation based on observations
  – no detailed representation of D or E
  – users must be identified
ACF: How it works

- Form virtual communities using clustering
  - e.g. k-Means clustering
  - requires a similarity (difference) metric,
    - e.g. mean squared difference
      \[
      \delta_{UJ} = \frac{1}{|InCommon|} \sum_{f \in InCommon} (U_f - J_f)^2
      \]
    - where Uf is Us rating of asset f
    - Pearson correlation coefficient may be better: \( r_{UJ} \)
      - (Shardanand & Maes 1995)
ACF: Lazy approach

- Build personalised community at run time:
  - Identify others with similarity above a threshold
  - Rating for an item is a weighted average of rating of similar users for that item.

\[
U_x = \bar{U} + \frac{\sum_{J \in \text{Raters of } x} (J_x - \bar{J})r_{UJ}}{\sum_{J \in \text{Raters of } x} |r_{UJ}|}
\]

Recommendations can be made without explicit rating:

Customers who bought this book also bought:

• Reinforcement Learning: An Introduction; R. S. Sutton, A. G. Barto

• Advances in Knowledge Discovery and Data Mining; U. M. Fayyad, et al

• Probabilistic Reasoning in Intelligent Systems; J. Pearl
ACF: Drawbacks

- No anonymity
- **The pump priming problem**: not enough observations ⇒ poor recommendations
PTV: The best of both worlds

(Smyth & Cotter, 1999)

- Generates personalised TV guides  
  - an intangible product
- Uses collaborative & case-based recommendations  
  - based on descriptions of programmes  
  - based on likes of users with similar tastes.
PTV: Recommendations & Feedback

Today's TV Listings

PTV by eMail

Grade Schedule Programme

Third Rock From The Sun
Alien Hunter Dick commands the rest of the family to throw him a surprise...adultery...

Sky 1
16:30-20:00

RTE1
20:00-21:00
Bread and Water Siobhan's husband Brendan with
the help of a cat. The programme moves with
the pace of a sprinter, and the dialogue with
ations can in a fit of rage. Light-hearted drama

RTE1
21:00-22:00

BBC 2
21:00-22:00

Hitchcock The Great
The first of a two-part 'Reputations special
detailing the life and work of Sir Alfred
Hitchcock. The most popular and most important
of all film directors. The programme follows
Hitchcock from the streets of so-called to
the movie spots of Hollywood and all points
between.

ER
21:30-22:15

Finale of Saving Grace and Gordon dance a
tango while Lucy pursues murder during the
psychiatric rotation. Medical drama starring
Anthony Edwards.
Represent or Observe?

Real World Scenario → Characterization → Representation Based

Observe

Historic Data → Recommend

Reasoning System → Solution

Representation Based
ACF v’s Feature Based

- ACF
  - requires identification
  - “representationless”
  - pump priming problem
  - scalability
  - problem with *consumable* assets

- Feature Based (CBR)
  - can be anonymous
    $\Rightarrow$ not personalised
  - requires representation
The Role for Machine Learning

- E-commerce systems should exploit available data on user behaviour and preferences.
- Learn from this data
  - This learning should be lazy because:
    - Incremental
    - Flexible
      - Lazy methods excel when the “use of experience” is not known a priori
        (Aha 1998)
e.g. Case-Based User Profiling

- User’s interaction with system forms a CB
  - e.g. Personal Travel Assistant in FIPA scenario
    - Foundation for Intelligent Physical Agents (www.fipa.org)
Case-Based PTA

• Help user plan and book travel:
  – Elaborate travel request based on profile
  – Select among offers from brokers
    (Waszkiewicz, Cunningham & Byrne, 1999)

• Travel behaviour stored as cases in a case-retrieval net.
  – Elaboration and selection tasks based on identifying similar cases in retrieval net.
What the PTA should learn

- You dislike connecting through Heathrow
- You like British Airways
- You will stay over a Saturday night
- You like a window seat
- You really dislike Heathrow
Elaborate Travel Request

Case Base

<table>
<thead>
<tr>
<th>Feature</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>OriginCountry</td>
<td>IR</td>
</tr>
<tr>
<td>OriginCity</td>
<td>Dublin</td>
</tr>
<tr>
<td>DestCountry</td>
<td>FR</td>
</tr>
<tr>
<td>DestCity</td>
<td>Paris</td>
</tr>
<tr>
<td>DepartureDate</td>
<td>990701</td>
</tr>
<tr>
<td>ReturnDate</td>
<td>990703</td>
</tr>
<tr>
<td>Company</td>
<td>null</td>
</tr>
<tr>
<td>Class</td>
<td>null</td>
</tr>
<tr>
<td>PreferredSeat</td>
<td>null</td>
</tr>
<tr>
<td>Meal</td>
<td>null</td>
</tr>
<tr>
<td>DepartureTimeAfter</td>
<td>null</td>
</tr>
<tr>
<td>DepartureTimeBefore</td>
<td>null</td>
</tr>
<tr>
<td>ReturnTimeAfter</td>
<td>null</td>
</tr>
<tr>
<td>ReturnTimeBefore</td>
<td>null</td>
</tr>
<tr>
<td>Via</td>
<td>null</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Feature</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>OriginCountry</td>
<td>IR</td>
</tr>
<tr>
<td>OriginCity</td>
<td>Dublin</td>
</tr>
<tr>
<td>DestCountry</td>
<td>FR</td>
</tr>
<tr>
<td>DestCity</td>
<td>Paris</td>
</tr>
<tr>
<td>DepartureDate</td>
<td>990701</td>
</tr>
<tr>
<td>ReturnDate</td>
<td>990703</td>
</tr>
<tr>
<td>Company</td>
<td>AirFrance</td>
</tr>
<tr>
<td>Class</td>
<td>Business</td>
</tr>
<tr>
<td>PreferredSeat</td>
<td>Window</td>
</tr>
<tr>
<td>Meal</td>
<td>Yes</td>
</tr>
<tr>
<td>DepartureTimeAfter</td>
<td>07.45</td>
</tr>
<tr>
<td>DepartureTimeBefore</td>
<td>09.45</td>
</tr>
<tr>
<td>ReturnTimeAfter</td>
<td>20.30</td>
</tr>
<tr>
<td>ReturnTimeBefore</td>
<td>22.30</td>
</tr>
<tr>
<td>Via</td>
<td>none</td>
</tr>
</tbody>
</table>

...
## Select Best Offer

<table>
<thead>
<tr>
<th>Feature</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>OriginCountry</td>
<td>IR</td>
</tr>
<tr>
<td>OriginCity</td>
<td>Dublin</td>
</tr>
<tr>
<td>DestCountry</td>
<td>FR</td>
</tr>
<tr>
<td>DestCity</td>
<td>Paris</td>
</tr>
<tr>
<td>DepartureDate</td>
<td>990701</td>
</tr>
<tr>
<td>ReturnDate</td>
<td>990703</td>
</tr>
<tr>
<td>Company</td>
<td>AirFrance</td>
</tr>
<tr>
<td>Class</td>
<td>Business</td>
</tr>
<tr>
<td>PreferredSeat</td>
<td>Window</td>
</tr>
<tr>
<td>Meal</td>
<td>Yes</td>
</tr>
<tr>
<td>DepartureTimeAfter</td>
<td>07.45</td>
</tr>
<tr>
<td>DepartureTimeBefore</td>
<td>09.45</td>
</tr>
<tr>
<td>ReturnTimeAfter</td>
<td>20.30</td>
</tr>
<tr>
<td>ReturnTimeBefore</td>
<td>22.30</td>
</tr>
<tr>
<td>Via</td>
<td>none</td>
</tr>
</tbody>
</table>

## Case Base

<table>
<thead>
<tr>
<th>Feature</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>OriginCountry</td>
<td>IR</td>
</tr>
<tr>
<td>OriginCity</td>
<td>Dublin</td>
</tr>
<tr>
<td>DestCountry</td>
<td>FR</td>
</tr>
<tr>
<td>DestCity</td>
<td>Paris</td>
</tr>
<tr>
<td>DepartureDate</td>
<td>990601</td>
</tr>
<tr>
<td>ReturnDate</td>
<td>990605</td>
</tr>
<tr>
<td>Company</td>
<td>AirFrance</td>
</tr>
<tr>
<td>Class</td>
<td>Business</td>
</tr>
<tr>
<td>Seat</td>
<td>Window</td>
</tr>
<tr>
<td>Meal</td>
<td>Yes</td>
</tr>
<tr>
<td>DepartureTime</td>
<td>08.45</td>
</tr>
<tr>
<td>ReturnTime</td>
<td>21.30</td>
</tr>
<tr>
<td>Via</td>
<td>none</td>
</tr>
<tr>
<td>Price</td>
<td>70</td>
</tr>
<tr>
<td>Currency</td>
<td>IRPound</td>
</tr>
</tbody>
</table>

| Accepted      | Yes         |

---

**ICCBBR August 1999**
Case-based User Profiling

- Uses data on customer to smooth interaction
- Requires identification
- Appropriate for business-to-business situations?
Conclusions

- Build effective recommender systems using case retrieval techniques
  - anonymous users
- ACF a powerful alternative
  - representationless
  - pump priming problem
- Use available data
  - e.g. case-based user profiling
References


References (cont’d)


