## Case-based Recommendation

Peter Brusilovsky with slides of Danielle Lee

## Where we are?

	Search	Navigation	Recommendation
Content-based			
Semantics / Metadata			
Social			

## Modern E-Commerce Site

Search CNET			views	cinet Cinet Rev
w To ~ Marketplace ~ Log In [	ad - CNETTV - H	News - Download -	Reviews -	
				NET > Reviews > Digital Cameras
		S	nera	<b>Digital Can</b>
See all digital cameras (	Find a digital can	Fi		— TOP DIGITAL CAMERA —
Manufacturer Zoom range Other	Price	Р		
Sony Less than 3X Flash memory	Less than \$100	Le		2 Nikon
Nikon 3X to 4X Digital camera	\$100 - \$200	\$1	1.01	Comments of the second
Panasonic 4X to 8X type	\$200 - \$300	\$2	<b>07000</b>	1
Canon 8X to 12X Resolution	\$300 - \$400	****	New York	Nikon D7000
Olympus More than 12X Maximum ISO	\$400 - \$500			Price: \$1,179.95 - \$2,179.95
Fujifilm See all zoom Weight	\$500 - \$1,200	\$5		
See all ranges Optical sensor type	See all prices	S	2	TOP DIGITAL CAMERA
See all type	See all prices	S		

## The Power of Metadata

- Modern e-commerce sites have a range of metadata for each item
  - Travel information presented in its price, duration, accommodation, location, mode of transport, etc.
  - Job information presented in the job kinds, salary, business category of each company, educational level, experience, location etc.
- This data is used in modern Faceted Search, more powerful than keyword search
- The power of metadata can be also used for better recommendation this is the essence of case-based way

#### **Digital cameras**

Flash me	morv Maximum IS	0 Weight	More			
SDXC Memo Card (7)		4 oz to 8 oz (7) More than 1 lb (1)	Optical sensor type Zoom range			
ou selected:						
Nikon 🚳 🚦	\$300 - \$400 🔯 12 mega	apixels 🚳 Compact	remove all 🚳			
results						
Show 10 \$	results per page Sort by	y: Review date 🛊	Compare Selecte			
	300 (Black)	\$329 at 2 stores				
Editors' rating:	little more creative control	Bottom Line: Snapshooters wanting a nore creative control than the average and-shoot offers should check out the Coolpix P300.				
	Specs: Digital camera - megapixels, AVCHD, CM lens - 4.3 mm - 17.9 mm Memory Card, SD Memor Memory Card, Pop-up fil sensor shift mechanism)	NOS, 4.2 x x Zoom - F/1.8-4.9, SDHC ory Card, SDXC ash, Optical (image				
	Price History	Set Price Alert	Note: The second			
	Nikon Coolpix SS Reviewed on 04/21/2011	9100 (Red)	\$326 to \$420 at 2 stores			
Editors' rating:	The Bottom Line: Nikon solid point-and-shoot cor with the Coolpix S9100 b performance, good photo feature-to-price ratio.	npact megazoom by offering fast	\$275 - eBay \$326 - Adorama			
	Specs: Digital camera - ( megapixels, H.264, CMC lens - 4.5 mm - 81 mm - Memory Card, SD Memor	DS, 18 x x Zoom F/3.5-5.9, SDHC				

### Metadata Could be Used in a Smarter Way

- "6 mega-pixel digital SLR for under \$200"
  - No result is returned → System slavishly respects customers' queries ("stonewalling")
- "Another camera like this one but with more optical zoom and a lower price"
  - Too complex for customers to provide this form of feedback directly to the system.
- "I never accepted the cameras above \$1000"
  - Few commercial system to remember customers' preferences over time.
  - Customers start their search from scratch in every visit.

## Case-Based Recommendation?

- A special form of *content-based* recommendation
- Assumes structured item information with a well defined set of features and feature values.
- Information are represented as a *case* and the system recommends the cases that are *most similar* to a user's preference
- Case-based representation also supports more advanced recommendation dialogues and explanations

## **Case-based Reasoning**

- Case-based recommendation origins in Case-Based Reasoning (CBR).
  - It is to solve new problems by reusing the solutions to problems that have been previously solved and stored as cases in a case-base.
  - Each case consists of a specification part, which describes the problem and a solution part, which describes the solution of the problem.
    - Solutions to similar prior problems are a useful starting point for new problem solving.
- "The users would like the similar one that they liked before."

### Simple Example of Case-based Recommendation Product #1

I want laptop having 250GB HDD, 1GB memory and 14 inch screen for \$400



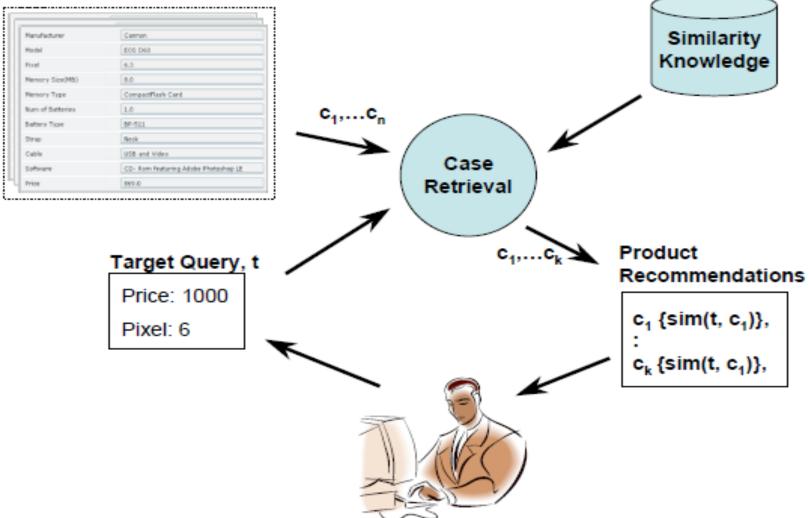
- •HDD: 250 GB
- Memory : 2 GB
- Screen Size : 15 inch
- •Price : \$550

Product #2 •HDD : 150 GB Memory : 1 GB Screen Size : 15 inch •Price : \$450

Product #3 •HDD: 250 GB Memory : 1 GB Screen Size : 14.2 inch •Price : \$420

### **Case-based Recommendation**

#### Product Case-Base

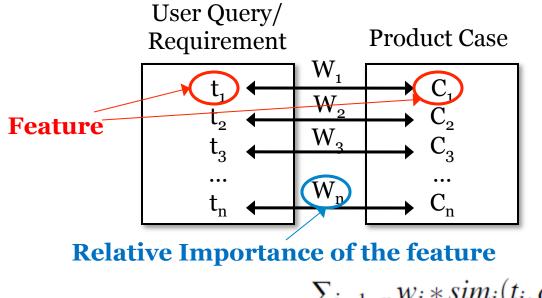


## **Case Representation**

	Manufacturer	Cannon			
Nominal Feature	Model	EOS D60			
reature	Pixel	6.3			
	Memory Size(MB)	8.0			
	Memory Type	CompactFlash Card			
	Num of Batteries	1.0			
	Battery Type	BP-511			
Numeric Feature	Strap	Neck			
reature	Cable	USB and Video			
	Software	CD- Rom featuring Adobe Photoshop LE			
	Price	869.0			

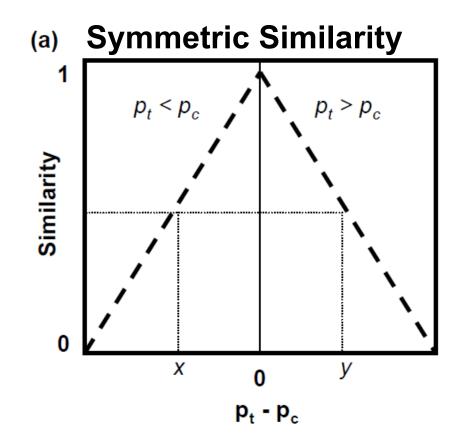
## Similarity Assessment (1)

• Similarity metrics that are based on an explicit mapping of case features and the availability of specialized feature level similarity knowledge.



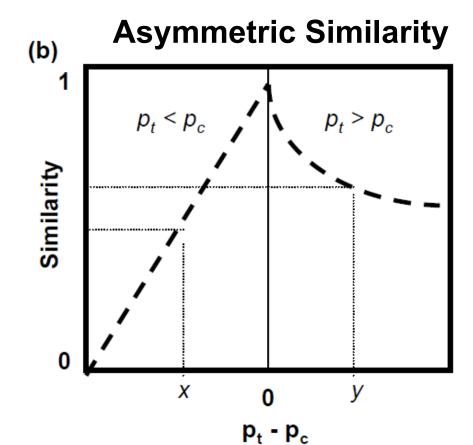
Similarity
$$(t,c) = \frac{\sum_{i=1..n} w_i * sim_i(t_i,c_i)}{\sum_{i=1..n} w_i}$$

## Similarity Assessment (2)



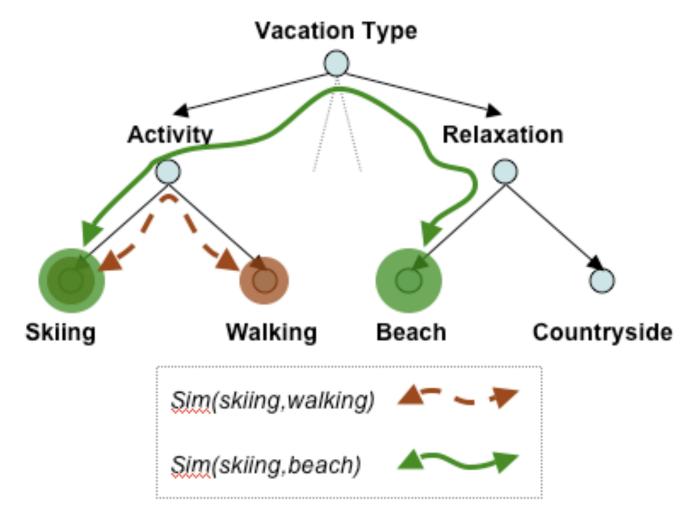
• In symmetric similarity, maximum similarity is achieved when a feature of a candidate case matches that of the target query. **No bias in favor** of either higher or lower values of the corresponding feature.

## Similarity Assessment (3)



• In asymmetric similarity, there is a **bias** to either higher or lower values (i.e. a product that is \$50 cheaper is better than \$50 more expensive)

### Similarity Assessment of Nominal Values



### **Partial Ontology of Vacation types**

Representing Similarity Knowledge

Danielle Lee

	Small Car	Medium Car	Large Car	SUV	Mini Van
SmCar	1	???	???	???	???
MdCar	???	1	???	???	???
LgCar	???	???	1	???	???
SUV	???	???	???	1	???
Minivan	???	???	???	???	1

15

### Acquiring Similarity Knowledge

- Based on knowledge made by a domain knowledge expert.
  - Normally it is hand-coded and expensive.
- Machine learning techniques.
  - Using several weight-learning algorithms, even knowledge-poor techniques can result in significant improvements in case-based classification tasks.
- Similarity assessment by users
  - A 'similarity teacher' evaluates the ordering for the given set of retrieval results.
  - The selections could be used not only for assessing the similarity but for acquiring users preference.

## **Case-based Job Recommendation**

Database Developer job for a financerelated company in Boston



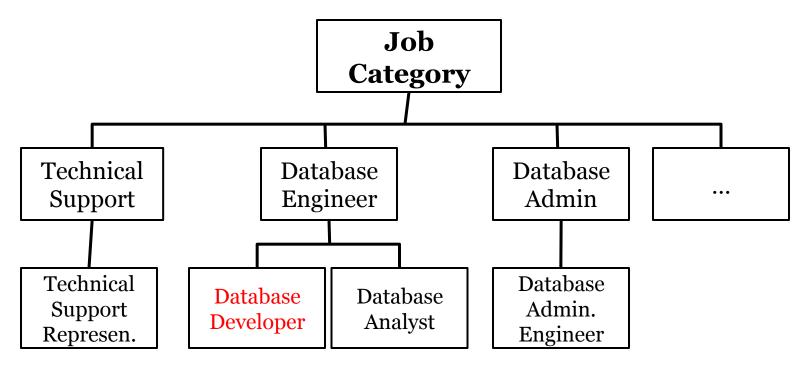
**Job #1** Database Analyst job for Company A

**Job #2** Database Administrator job for Company B

**Job #3** Technical Support Engineer for Company C

## Job Related Knowledge (1)

• Partial Ontology about job category.



# Job Related Knowledge (2)

- Taxonomy about Company
  - Company A : Insurance company, downtown in Boston.
  - Company B : Pharmaceutical company, 5 miles distance from Boston.
  - Company C : Computer manufacturing domain,
    1.5 miles distance from Boston.

### **Proactive** - Job Recommendation System

#### Proactive

#### [Preference] [Feedback] [Help] [Logout]

:: Matched jo	bs to your preference within 3d	ays ::								
Total: 50 (1/3)	)	≷ Saved Job 🐐	Recommene	ded Job	Visualize these li	st		I	Preferen	ce] 1 2 3
Job Category	Title	Company	City	State	Position Type	Salary	Experience	Education	Post Date	Relevance
Database Manager	Marketing Database Manager [SAVE]	Info Technologies, Inc.	New York	NY	Full-Time, Contract	unspecified	5-10 Years Experience	unspecified	01-23-08	****
Database Architect	3 NEW TECH OPPORTUNITIES! - .NET DVELOPER (C++, C#), DATABASE ARCHITECT, .NET DEVELOPER MANAGER [SAVE]	Concepts in Staffing	New York	NY	Full-Time, Employee	unspecified	5-10 Years Experience	Bachelor's degree	01-23-08	****
Database Architect	Database Architect [SAVE]	Spot Runner	Los Angeles	CA	Full-Time, Employee	unspecified	10-15 Years Experience	Bachelor's degree	01-23-08	****
Database Architect	Solutions Architect   Senior Solutions Architect [SAVE]	Silver Key	San Francisco	CA	Full-Time, Employee	unspecified	2-5 Years Experience	Bachelor of Science	01-23-08	****
Database Architect	ETL Architect [SAVE]	Kinetic Networks, Inc.	San Francisco	CA	Full-Time, Employee	unspecified	2-5 Years Experience	unspecified	01-23-08	****
Database Specialist	EDUCATIONAL SALES SPECIALIST - EAST [SAVE]	Sika Sarnafil, Inc.	New York	NY	Full-Time, Employee	unspecified	5-10 Years Experience	Bachelor of Science	01-23-08	****
Database Architect	MySQL DBA / Architect [SAVE]	Slide, Inc	San Francisco	CA	Full-Time, Employee	unspecified	2-5 Years Experience	unspecified	01-23-08	****
Database Architect	SQL DAtabase Architect [SAVE]	Global Technical Talent	New York City	NY	Full-Time, Contract	unspecified	2-5 Years Experience	unspecified	01-23-08	****
Database Manager	Release Manager [SAVE]	Kaiser Permanente	Oakland	CA	Full-Time, Employee	unspecified	5-10 Years Experience	unspecified	01-23-08	***
Database Architect	Sr. Database Architect [SAVE]	Careers on the Move	New York	NY	Full-Time, Employee	\$130,000 to \$150,000 per year	10-15 Years Experience	Master of Science	01-23-08	***
Database Developer	Java Developer - Database Developer [SAVE]	Adobe Systems	San Francisco	CA	Full-Time, Employee	unspecified	5-10 Years Experience	Bachelor of Science	01-21-08	***
Technical Support Consultant	Patient Monitoring Technical Consultant [SAVE]	Philips North America	New York	NY	Full-Time, Employee	unspecified	2-5 Years Experience	unspecified	01-20-08	***
Technical Support Consultant	Summit System Support Consultant [SAVE]	Solomon-Page Group	New York	NY	Full-Time, Contract	\$450 to \$650 per year	5-10 Years Experience	Bachelor's degree	01-20-08	***

### Diversity

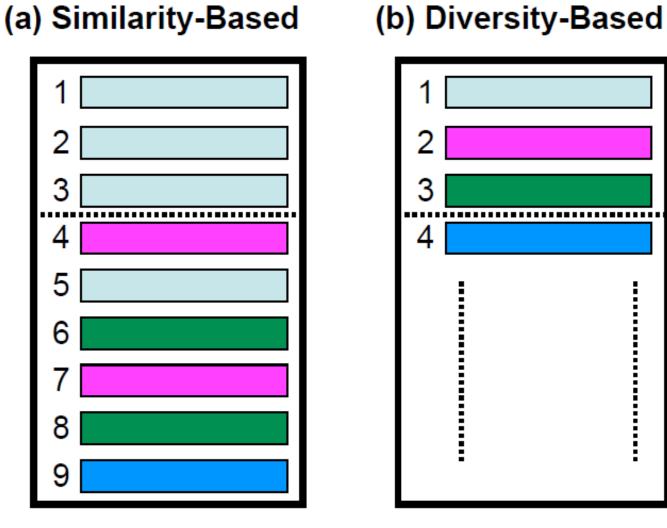
"I want a 2-week vacation for two in the sun, costing less than \$750, within 3 hours flying time of Ireland. I expect good night-life and recreation facilities on-site"

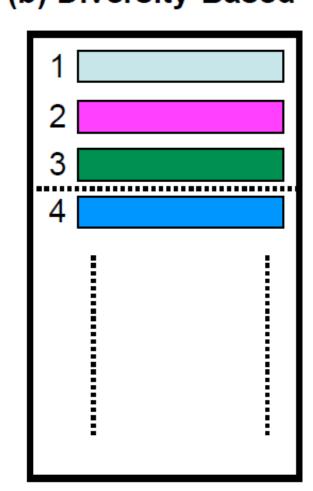


### System suggests ...

- 1. Hercules Complex in the Costa Del Sol, Spain on the first two weeks of July
- 2. Hercules Complex in the Costa Del Sol, Spain on the first two weeks of August
- 3. Pleasure Complex in the Costa Del Sol, Spain on the last two weeks of July
- 4. Hercules Complex in the Costa Del Sol, Spain on the last two weeks of July
- 5. ...

## Similarity vs. Diversity (1)





# Similarity vs. Diversity (2)

- Bounded Random Selection: from the top *bk* most similar cases to the target query, select *k* random cases.
  - The diversity could increase but the similarity could also decrease.
- Bounded Greedy Selection: define the diversity of a set of retrieved cases to be the average dissimilarity between all pairs of these cases.
  - 50% improvement in relative diversity with a minor loss of less than 10% in similarity to the target query.
  - A unit drop in similarity can be traded for almost 3 units of diversity using this method.
  - Increased computational efficiency.

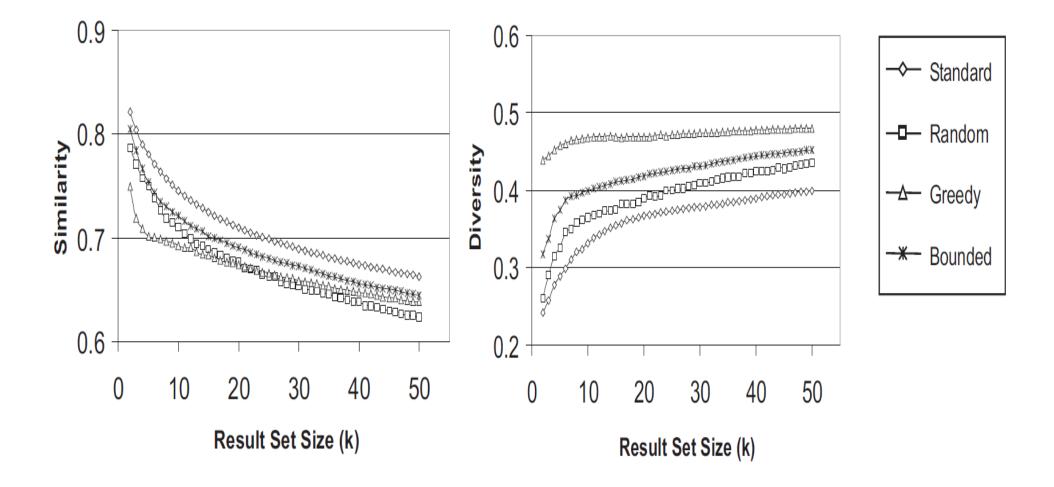
## **Bounded Greedy Selection**

- Proposed in Smyth & McClave (2001).
- The key idea is the quality metric that combines diversity and similarity.
  - 1. Select the best *bk* cases according to the similarity.
  - 2. Pick up the one with the highest similarity.
  - 3. During each subsequent iteration, the case with the highest combination of similarity and diversity with respect to the set of cases selected during the previous iteration.

$$Quality(t, c, R) = Similarity(t, c) * RelDiversity(c, R)$$
 (11.4)

$$RelDiversity(c,R) = 1 \text{ if } R = \{\}; \\ = \frac{\sum_{i=1..m} (1 - Similarity(c,r_i))}{m} \text{, otherwise}$$
(11.5)

## Similarity vs. Diversity (3)



# Other approaches for diversity

- Shimazu (2001, 2002)
  - It is suitable to the situations where the set of recommended cases is drawn from a set of sufficiently similar cases.
- Similarity layers
  - A set of cases are ranked by their similarity to the target query and partitioned into similarity layers.
  - The lowest similarity layer is selected using an optimal diversity maximizing algorithm.
- Similarity intervals
- Order-based Retrieval
- Compromise-driven approach

# **Conversational Recommendation**

- Many recommenders operate in a reactive and *single-shot fashion*.
  - If users don't find what they want, the only option is to revise their query and start again.
  - User's initial query could be adapted and refine as they learn more about a particular product-space.
- Conversational recommender is to engage users in an extended, interactive recommendation dialog during the query elicitation.
  - Navigation by asking
  - Navigation by proposing

### Navigation by asking (1) - Adaptive Place Advisor

- 1. Inquirer: Where do you think I should eat tonight?
- 2. Advisor: What type of food would you like?
- 3. Inquirer: What types are there?
- 4. Advisor: You can say things like Chinese, Indian, and Mediterranean.
- 5. Inquirer: Oh, maybe a cheap Indian place.
- Advisor: What kind of parking w ould you like?
- 7. Inquirer: I don't care, as long as it's in Palo Alto.
- 8. Advisor: I'm sorry, I don't know of any restaurants like that, would you like to search for any pirice?
- 9. Inquirer: No, I think I'd like Chinese instead.
- 10. Advisor: How does Mandarin Gourmet on 420 Ramona sound?1
- 11. Inquirer: No, what else do you have?
- 12. Advisor: How about Jing-Jing Szechwan Hunan Gourmet on 443 Emerson Street?
- 13. Inquirer: Sure, that sounds fine.

# Navigation by asking (2)

- The selection of questions does matter.
  - Enthropy -Dolye & Cunningham (2000)
  - Information gain Quinlan (1986); Quinlan (1993) and Shimazu (2001); Shimazu (2003)
  - Influence Schmitt (2002)
- However, directly asking questions to users can be burden to them.
  - A long list of questions, uncertain answers or rejected questions.
  - Significant interfacing burden

# Navigation by proposing (1)

- The users are presented with one of more recommendation alternatives, rather than a question, during each recommendation cycle.
  - Rating based feedback
  - Critique based feedback: Constraints over certain features of recommendations
  - Preference based feedback: Expressed preference for one alternative over the others

## Critique based feedback

UKRAINIAN VILLAGE. TWO bedroom rehab garden apartment. Lr. Eurokitchen, hwfl, excellent security, forced air, lots of closets, laundry in building. Garage space included. Dogs OK. Available immediately. \$600/ mo. 312-489-1554./ ; 60622 Phone: 312-489-1554 2-bedrooms \$600 (West Town Bucktown) This apartment is OK, but make it ... bigger cheaper nicer safer This neighborhood could be more ... convenient conservative dynamic

# **Compound Critiques**

QUIKEHOR	COM	. F			HOME : ABOUT THIS PROJECT : CONTACT
>> Digital Cameras			Unit Critiqu	ues	
Shop for: ▶ Digital Cameras ▶ Co	-				
	A djust your prefe		to find the right ca	mera tor you	Explain:
Canon	Manufacturer	Х	Canon	x	1. Less Memory and Lower Resolution and Cheaper
	Optical Zoom	Ŧ	7x	<b>1</b>	This Critique covers <b>153</b> other
	Memory (MB)	Ŧ	512	<b>1</b>	Digital Cameras
	Weight (Grams)	Ŧ	780	Ť	Less Memory Current Value: 512 MB
Page 1994	Resolution	Ŧ	6.2 M Pixels	Ť	Critique: Less Than Remaing: (0 to 256 MB)
Product Found: Canon EOS 30	Size	X	Large	X	Lower Resolution
6.3 Megapixel CMOS sensor 7-point wide-area AF High-performance DIGIC processor	Case	Х	Magnesium	X	Current Value: 6.2 M Pixels Critique: Less Than
100-1600 ISO speed range Compatible with all Canon EF	Price	Ŧ	995	1	Remaing: (1.4 to 5.9 M Pixels)
lenses and EX Speedlites PictBridge, Canon Direct Print and Bubble let Direct econoctible	We have more mat	ching	cameras with the follo	owing:	Cheaper Current Value: 995 €
Bubble Jet Direct compatible – no PC required	1. Less Memory and L	ower R	esolution and Cheaper	EXPLAIN PICK	Critique: Less Than Remaing: (75€to 960€)
I've found the Camera I want! 🖒	2. Different Manufactu	rer and	Less Zoom and Lighter	EXPLAIN PICK	PICK
No lets start again 🔗	3. Lighter and Smaller	and Dif	ferent Case	EXPLAIN PICK	

## Explanations and Clustering (Pu)

The top car	ndidate	according	<mark>, to you</mark>	r preferences					
Manufacturer	Price	MegaPixels	Optical zoom	Memory type	Flash memory	LCD screen size	Depth	Weight	
Canon	\$242.00	5.0 MP	Зх	CompactFlash Card	32 MB	1.8 in	1.37 in	8.3 oz	<u>choose</u>

We have n	We have more products with the following										
they are ch	eaper and	d lighter, b	ut hav	e fewer megapix	æls						
Nikon	\$167.95	4 MP	Зх	SD Memory Card	14 MB	1.8 in	1.4 in	4.6 oz	<u>choose</u>		
Canon	\$230.00	4.1 MP	Зх	CompactFlash Card	32 MB	1.5 in	1.09 in	6.53 oz	<u>choose</u>		
Canon	\$180.00	3.3 MP	Зх	SD Memory Card	16 MB	2 in	0.83 in	4.06 oz	<u>choose</u>		
Canon	\$219.18	4.2 MP	4x	MultiMedia Card	16 MB	1.8 in	1.51 in	6.35 oz	<u>choose</u>		
Canon	\$163.50	3.2 MP	4x	MultiMedia Card	16 MB	1.8 in	1.5 in	6.3 oz	<u>choose</u>		
Canon	\$199.40	3.2 MP	2.2x	SD Memory Card	16 MB	1.5 in	1.4 in	5.8 oz	<u>choose</u>		

they have n	nore meg	apixels and	l bigger	screens, but a	re more ex	pensive	;		
Sony	\$365.00	7.2 MP	Зх	Internal Memory	32 MB	2.5 in	1.5 in	6.9 oz	<u>choose</u>
Canon	\$439.99	7.1 MP	Зх	SD Memory Card	32 MB	2 in	1.04 in	6 oz	<u>choose</u>
Fuji	\$253.00	6.3 MP	4x	XD-Picture Card	16 MB	2 in	1.4 in	7.1 oz	<u>choose</u>
Sony	\$336.00	7.2 MP	Зх	Internal Memory	32 MB	2 in	1 in	5 oz	<u>choose</u>
Nikon	\$304.18	7.1 MP	Зх	Internal Memory	13.5 MB	2 in	1.4 in	5.3 oz	<u>choose</u>
Olympus	\$334.00	7.4 MP	5x	XD-Picture Card	32 MB	2.0 in	1.7 in	7.1 oz	<u>choose</u>

they are lig	hter and	thinner, bu	t have l	less flash mem	огу				
Pentax	\$238.99	5.3 MP	Зх	Internal Memory	10 MB	1.8 in	0.8 in	3.7 oz	choose
Canon	\$273.18	4.0 MP	Зх	SD Memory Card	16 MB	2 in	0.82 in	4.59 oz	choose
Nikon	\$329.95	5.1 MP	Зх	Internal Memory	12 MB	2.5 in	0.8 in	4.2 oz	choose
Canon	\$316.18	5.3 MP	Зх	SD Memory Card	16 MB	2 in	0.81 in	4.59 oz	choose
Casio	\$386.00	7.2 MP	Зх	Internal Memory	8.3 MB	2.5 in	0.88 in	4.48 oz	choose
Fuji	\$309.18	6.3 MP	Зх	XD-Picture Card	16 MB	2.5 in	1.1 in	5.5 oz	choose

# Case-based User Profiling (1)

- Conversational recommenders can react to the feedback provided by users within each session.
  - In-session personalization only two users who respond in the same way within a session will receive the same recommendations.
  - How can the systems adapt to the users' persistent preference?
- It is important for the recommenders to learn and maintain a long-term model of a user's recommendation preferences.

# Case-based User Profiling (2)

- CB leverages available content descriptions of cases as a form of case-based user profile.
  - User profile is made of a set of cases and the preference (like or dislike)
- CASPER : Online recruitment system using implicit user profile (from positive and negative points of view) and this profile is used to re-order the recommendations.
  - The Personal Travel Assistant also has similar approach.

## Feature Level User Profiling

- The preference related to features and their values such as preferred values for a particular features, the relative importance of a particular attributes, etc.
  - In restaurant recommendation, the kind of cuisine has an importance weight of 0.4 and parking facilities have a preference weight of 0.1. The user also prefers Italian cuisine with 0.35 weight to German food with 0.1 weight.

# Hybridization of CB and CF - PTV

- O' Sullivan, et al. (2002)
- To solve sparsity problem or latency problem in CF, case-based technology was used.
- By the derived similarity knowledge using data mining technology, the relationships between information items was extended.
- Increased recommendation coverage and recommendation accuracy.

# Question?