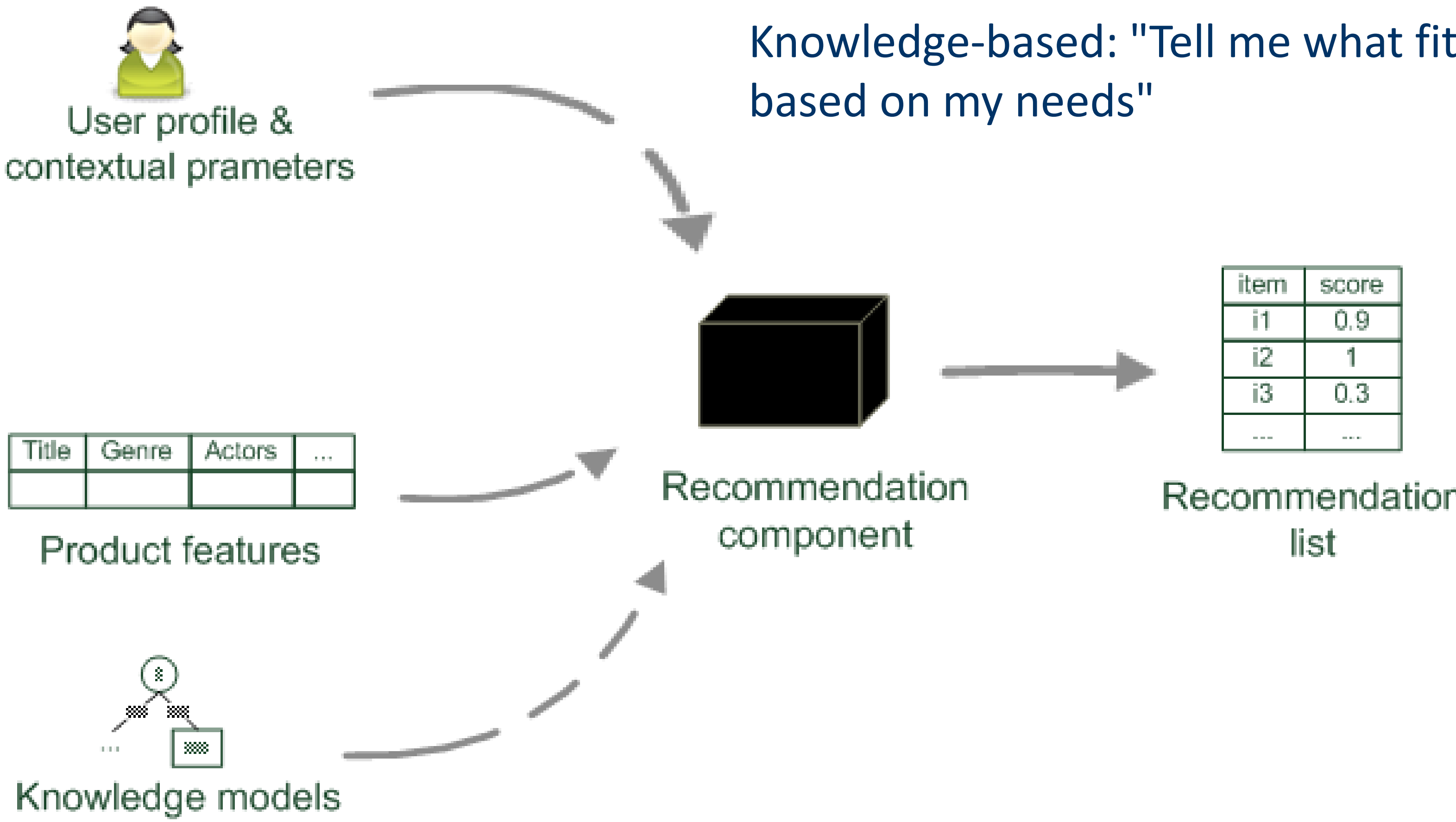
Knowledge-based recommenders

Nisheeth

Knowledge-based recommendation

Knowledge-based: "Tell me what fits based on my needs"



Knowledge-based recommendation

- Explicit domain knowledge
  - Sales knowledge elicitation from domain experts
  - System mimics the behavior of experienced sales assistant
  - Best-practice sales interactions
  - Can guarantee "correct" recommendations (determinism) with respect to expert knowledge

- Conversational interaction strategy
  - Opposed to one-shot interaction
  - Elicitation of user requirements
  - Transfer of product knowledge ("educating users")

Types

- Different views on “knowledge”
  - Logic-based knowledge descriptions (from domain expert)
    - E.g. Hard and soft constraints
  - Utility-based RS
    - E.g. MAUT – Multi-attribute utility theory

Logic-based knowledge base

- Design an RS knowledge base
  - Customer properties (V_C)
  - Product properties (V_{PROD})
  - Fundamental domain constraints (C_R)
  - Optional (filter) constraints (C_F)
  - Input requirements (C_C)
- Useful to represent using first order logic
  - Represent products as conjunctions of features
- Can treat RS as a constraint satisfaction problem

$$CSP(V_c, V_{PROD}, C_R \cup \ C_F \cup \ C_{PROD} \ \cup C_C)$$

- Goal: Output some logically <u>consistent</u> V_{PROD} from $C_R \cup \ C_F \cup \ C_{PROD} \ \cup C_C$

Logical consistency check

- A set of statements is logically consistent if they can all be simultaneously true
- Shall we work through some examples?
- I am a man. I have short hair. You have long hair. You are a woman.
- Everyone should be tolerant because there is no way to judge another person's beliefs.
- It is raining. It is not raining.
- Light is simultaneously both a wave and a particle.
- God can do anything.
- This sentence is false.

Typical solution approaches

- Backtracking
  - Recursive depth first search
- Constraint propagation (e.g. AC-3)
  - Store arcs that represent constraints between variable pairs
  - Eliminate one variables possible values based on constraints
  - Iterate
- Local search (e.g. min-conflicts)
  - Assign values to all variables
  - Pick a violating variable
  - Assign a value that minimizes conflicts for it
  - Iterate
- Historically computationally complex, but recent work shows promise of scalability

Not well suited for

- Situations with subjective preferences
- Because of conjunction fallacies
- *Sharmishtha is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in antiwar demonstrations*.
- Assign probabilities to the following sentences being true
  - She is an investment banker.
  - She is active in the feminist movement.
  - She is an investment banker and is active in the feminist movement.

Sample constraint-based problem

$V_C = \{kl_c$: [expert, average, beginner] . . . . /* level of expertise */
wr_c: [low, medium, high] . . . . /* willingness to take risks */
id_c: [shortterm, mediumterm, longterm] . . . . /* duration of investment */
aw_c: [yes, no] . . . . /* advisory wanted ? */
ds_c: [savings, bonds, stockfunds, singleshares] . . . . . . /* direct product search */
sl_c: [savings, bonds] . . . . /* type of low-risk investment */
av_c: [yes, no] . . . . /* availability of funds */
sh_c: [stockfunds, singlshares] . . . . /* type of high-risk investment */ $\}$

V_{PROD} = {$name_p$: [text] .... /* name of the product */
er_p: [1..40] .... /* expected return rate */
ri_p: [low, medium, high] .... /* risk level */
$mniv_p$: [1..14] .... /* minimum investment period of product in years */
$inst_p$: [text] .... /* financial institute */ }

$$C_R = \{CR_1: wr_c = high \rightarrow id_c \neq shortterm,$$
$$CR_2: kl_c = beginner \rightarrow wr_c \neq high\}$$

$$C_F = \{CF_1: id_c = shortterm \rightarrow mniv_p < 3,$$
$$CF_2: id_c = mediumterm \rightarrow mniv_p \geq 3 \wedge mniv_p < 6,$$
$$CF_3: id_c = longterm \rightarrow mniv_p \geq 6,$$
$$CF_4: wr_c = low \rightarrow ri_p = low,$$
$$CF_5: wr_c = medium \rightarrow ri_p = low \vee ri_p = medium,$$
$$CF_6: wr_c = high \rightarrow ri_p = low \vee ri_p = medium \vee ri_p = high,$$
$$CF_7: kl_c = beginner \rightarrow ri_p \neq high,$$
$$CF_8: sl_c = savings \rightarrow name_p = savings,$$
$$CF_9: sl_c = bonds \rightarrow name_p = bonds\ \}$$

Logical product description

$$C_{PROD} = \{CPROD_1: name_p = savings \wedge er_p = 3 \wedge ri_p = low \wedge mniv_p = 1 \wedge inst_p = A;$$
$$CPROD_2: name_p = bonds \wedge er_p = 5 \wedge ri_p = medium \wedge mniv_p = 5 \wedge inst_p = B;$$
$$CPROD_3: name_p = equity \wedge er_p = 9 \wedge ri_p = high \wedge mniv_p = 10 \wedge inst_p = B\}$$

$$C_C = \{wr_c = low, kl_c = beginner, id_c = shortterm, sl_c = savings\}$$

What will the RS output be?

Query relaxation

- What if no exact match is found?
- Approach: find maximal subset of query that removes conflict
- Naïve solution has exponential complexity
- Can you design a better solution?

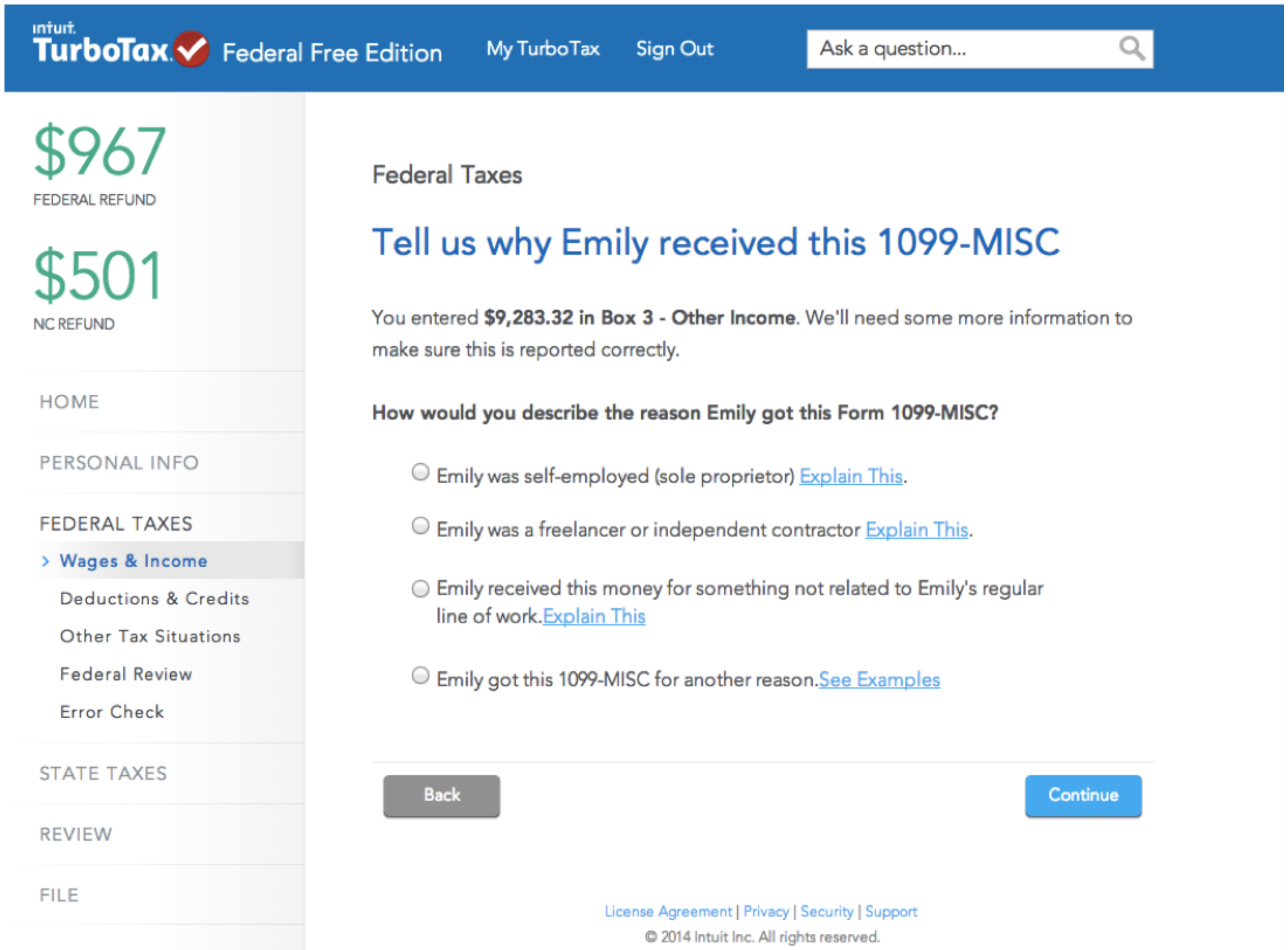
ID	Product p1	Product p2	Product p3	Product p4
CF_1	0	1	0	1
CF_2	1	0	1	0
CF_3	0	1	1	0
CF_4	1	1	0	1

Iterate over products to find # relaxations needed to satisfy

Can also involve customer

- Computation of minimal revisions of requirements
  - Do you want to relax your brand preference?
    - Accept *Panasonic* instead of *Canon* brand
  - Or is photographing landscapes with a wide-angle lens and maximum cost less important?
    - Lower focal length > 28mm and Price > 350 EUR
  - Optionally guided by some predefined weights or past community behavior
- Be aware of possible biases (e.g. age, family status, …)

Use case


intuit.
TurboTax
Federal Free Edition
My TurboTax
Sign Out
Ask a question...
$967
FEDERAL REFUND
$501
NC REFUND
HOME
PERSONAL INFO
FEDERAL TAXES
Wages & Income
Deductions & Credits
Other Tax Situations
Federal Review
Error Check
STATE TAXES
REVIEW
FILE
Federal Taxes
Tell us why Emily received this 1099-MISC
You entered $9,283.32 in Box 3 - Other Income. We'll need some more information to make sure this is reported correctly.
How would you describe the reason Emily got this Form 1099-MISC?
Emily was self-employed (sole proprietor) Explain This.
Emily was a freelancer or independent contractor Explain This.
Emily received this money for something not related to Emily's regular line of work. Explain This
Emily got this 1099-MISC for another reason. See Examples
Back
Continue
License Agreement | Privacy | Security | Support
© 2014 Intuit Inc. All rights reserved.

Best suited for

	Low	Medium	High
Upfront cost			X
Error cost			X
Heterogeneity	X		
Frequency	X		
Scale			X

Utility-based knowledge bases

- Multi-Attribute Utility Theory (MAUT)
  - Each item is evaluated according to a predefined set of dimensions that provide an aggregated view on the basic item properties
- E.g. quality and economy are dimensions in the domain of digital cameras

id	value	quality	economy
price	≤250	5	10
	>250	10	5
mpix	≤8	4	10
	>8	10	6
opt-zoom	≤9	6	9
	>9	10	6
…	…	…	…

Customer-item utilities with MAUT

- Customer interests:

customer	quality	economy
Cu_1	80%	20%
Cu_2	40%	60%

*

- Item utilities:

quality	economy	utility: cu_1	utility: cu_2
P1 Σ(5,4,6,6,3,7,10) = 41	Σ (10,10,9,10,10,10,6) = 65	45.8 [8]	55.4 [6]
P2 Σ(5,4,6,6,10,10,8) = 49	Σ (10,10,9,10,7,8,10) = 64	52.0 [7]	58.0 [1]
P3 Σ(5,4,10,6,10,10,8) = 53	Σ (10,10,6,10,7,8,10) = 61	54.6 [5]	57.8 [2]
...	...	...	...

**

$$utility(p) = \sum_{j=1}^{\#(dimensions)} \underbrace{interest(j)}_{*} * \underbrace{contribution(p,j)}_{**}$$

Output items ranked by utility

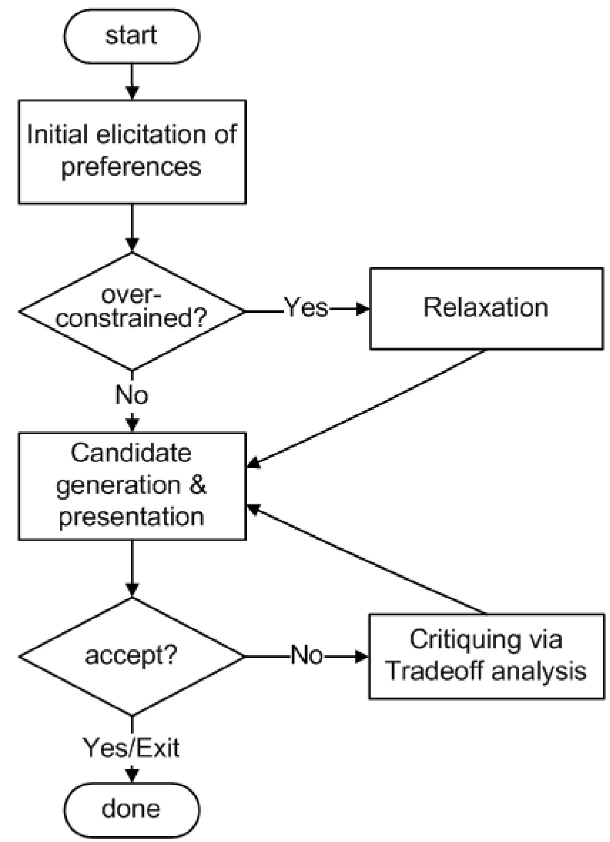
Constraint-based recommendation III

- More variants of recommendation task
  - Customers maybe not know what they are seeking
  - Find "diverse" sets of items
    - Notion of similarity/dissimilarity
    - Idea that users navigate a product space
    - If recommendations are more diverse than users can navigate via critiques on recommended "entry points" more efficiently (less steps of interaction)

  - Bundling of recommendations
    - Find item bundles that match together according to some knowledge
      - E.g. travel packages, skin care treatments or financial portfolios
      - RS for different item categories, CSP restricts configuring of bundles

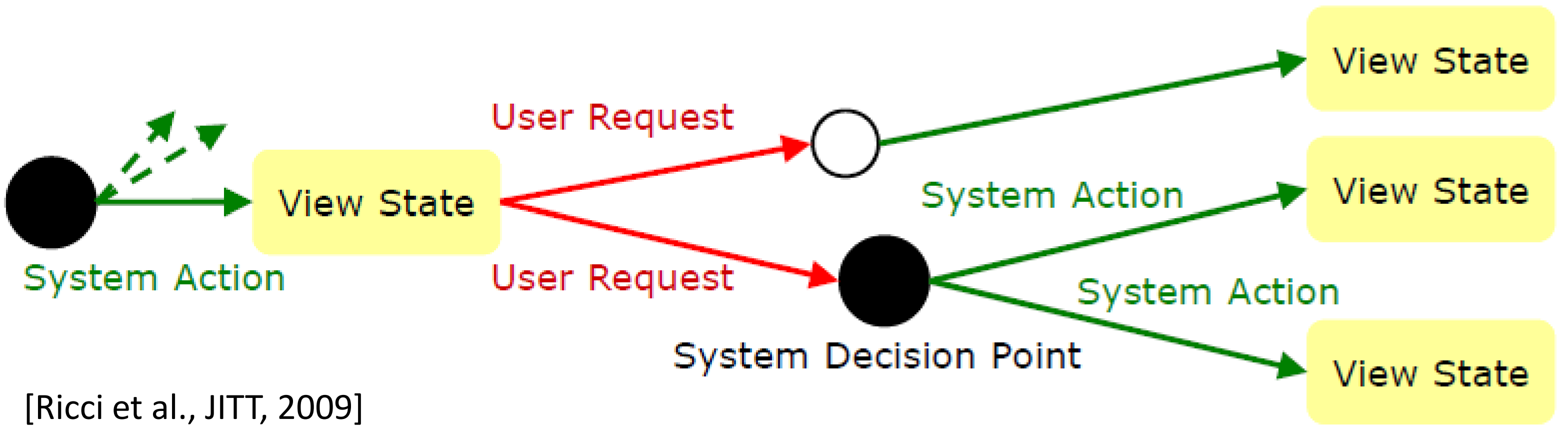
Conversational strategies

- Process consisting of multiple conversational moves
  - Resembles natural sales interactions
  - Not all user requirements known beforehand
  - Customers are rarely satisfied with the initial recommendations
- Different styles of preference elicitation:
  - Free text query interface
  - Asking technical/generic properties
  - Images / inspiration
  - Proposing and Critiquing



Example: adaptive strategy selection

- State model, different actions possible
  - Propose item, ask user, relax/tighten result set,...



[Ricci et al., JITT, 2009]

Limitations of knowledge-based recommendation methods

- Cost of knowledge acquisition
  - From domain experts
  - From users
  - Remedy: exploit web resources

- Accuracy of preference models
  - Very fine granular preference models require many interaction cycles with the user or sufficient detailed data about the user
  - Remedy: use collaborative filtering, estimates the preference of a user

    However: preference models may be instable

    - E.g. asymmetric dominance effects, conjunction effects, sunk cost effects