Optimizing Web Shop Facet Navigation

Damir Vandic, Flavius Frasincar, Uzay Kaymak

vandic@ese.eur.nl
Erasmus University Rotterdam

Amazon.com: Cell Phones

www.amazon.com/s/ref=lp_2407749011_nr_p_72_0?rh=n%3A2335752011%2Cn%3A%212335753011%2Cn%3A7(

Reader

Apple Wikipedia Popular Tweakers

amazon

Try Prime

1-24 of 2,353 results for **Cell Phones & Accessories : Cell Phones**

Sort by New and Popular

Refine by

Featured Brands

- Samsung (453)
- Motorola (114)
- Nokia (239)
- Apple (136)
- LG (175)
- See more

Display Size

- 3.9 Inches & Under (481)
- 4.0 to 4.4 Inches (346)
- 4.5 to 4.9 Inches (228)
- 5.0 to 5.4 Inches (258)
- 5.5 Inches & Over (160)

Color

Operating System

- Android (1,046)
- BlackBerry (126)
- iOS (156)
- Windows Phone (132)

Internal Storage

- 4 GB (145)
- 8 GB (232)

See Color Options

Samsung GT-i8190 Galaxy S3 Mini 3G GSM

~~$599.99~~ **$195.53**

In stock on July 20, 2014

FREE Shipping

More Buying Choices

$195.53 new (23 offers)

$179.00 used (14 offers)

(1,266)

See Color Options

BLU Advance 4.0 Unlocked Dual SIM Phone (White)

~~$99.00~~ **$85.21**

Order in the next **33 hours** and get it by **Monday, Jul 21.**

FREE Shipping

More Buying Choices

$79.99 new (19 offers)

$80.00 used (1 offer)

(235)

See Color Options

Samsung Galaxy S4 Mini GT-I9192 GSM Factory Unlocked Dual Sim - White

~~$699.99~~ **$292.53**

In stock on July 20, 2014

FREE Shipping

More Buying Choices

$292.53 new (22 offers)

$299.00 used (3 offers)

(243)

Samsung Galaxy S5 SM-G900H 16GB Factory Unlocked International Version

Samsung Galaxy S3 GT-i8190 Mini Blue 8GB factory Unlocked 3G

See Size Options

Motorola Moto G - Global GSM - Unlocked - 16GB (Black)

Introduction

- Terminology:
  - properties (e.g., `Color`)
  - values (e.g., `Red`)
  - facets, e.g.,
    - `Color:Red`
    - `Color:Red`
    - `WiFi:true`
    - `Price:64.00`

Introduction

Assumptions

- Facets:
  - Qualitative (nominal, boolean)
  - Numeric (integer, double)
- Numeric facets treated differently

Introduction

- Faceted search vs keyword-based search
  - browsing
  - progressive query refinements

Introduction

- Faceted search vs keyword-based search
  - browsing
  - progressive query refinements
- Open issues
  - too many facets to be shown at once
  - usually fixed, manually curated, facet list
  - not optimal due to changing queries

Introduction

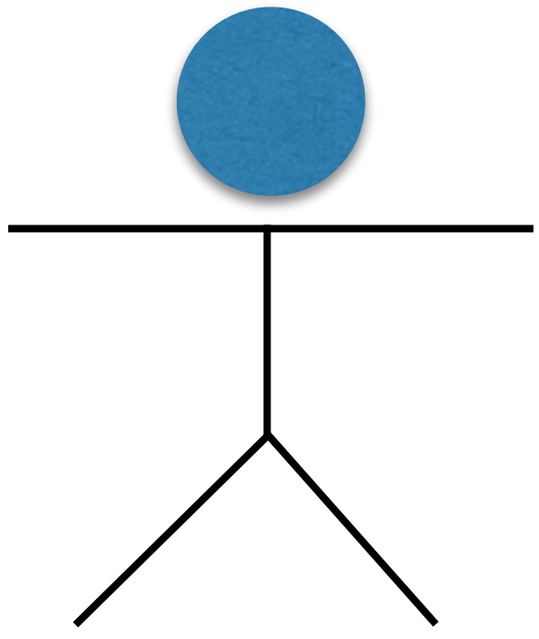
Our approach

- ranks properties and their corresponding
- query dependent
- employs impurity measures
- weighting scheme for coverage bias

Approach

Intuition

Bob



Brand:Samsung

Audio Formats:MP3

Approach

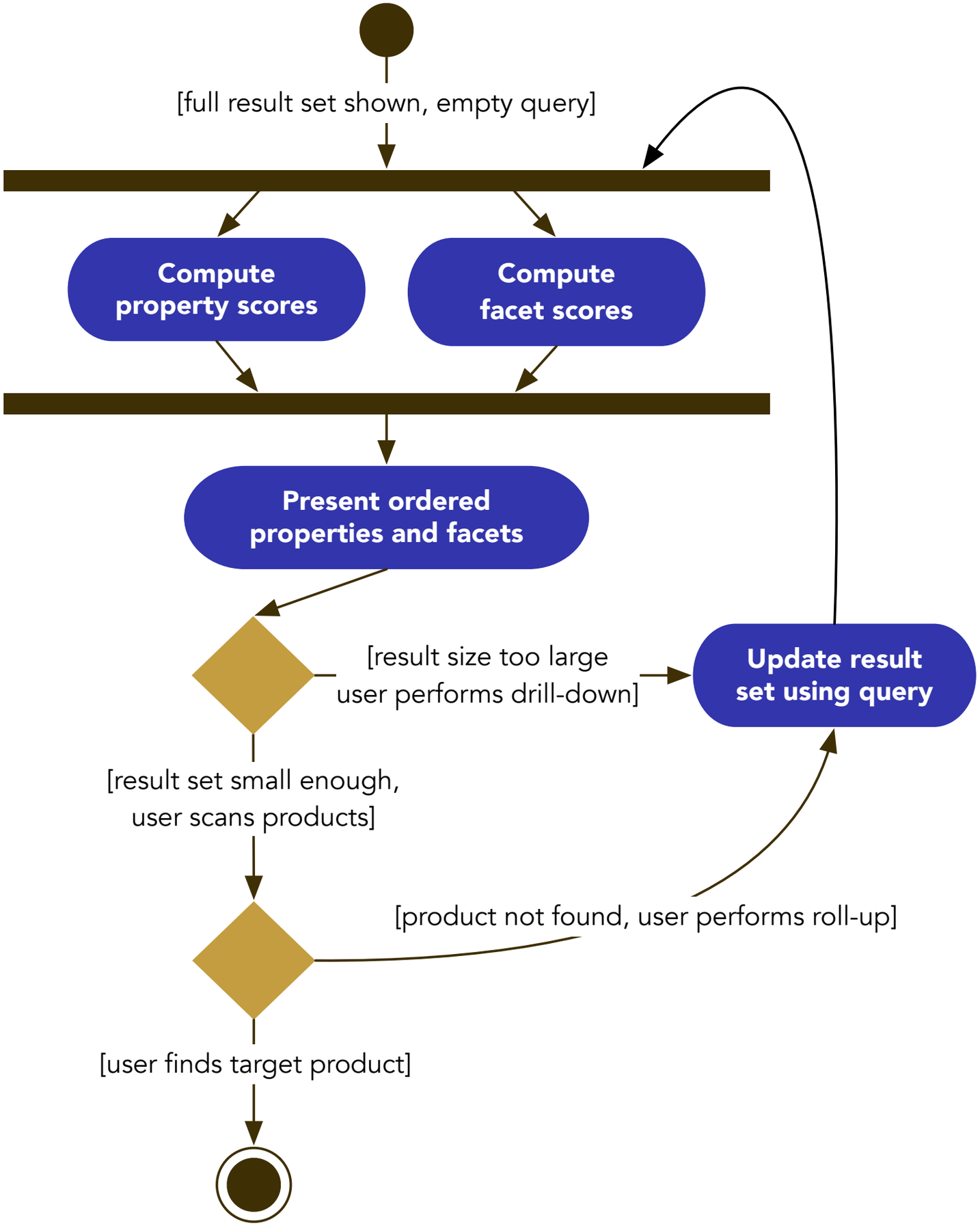
Search sessions

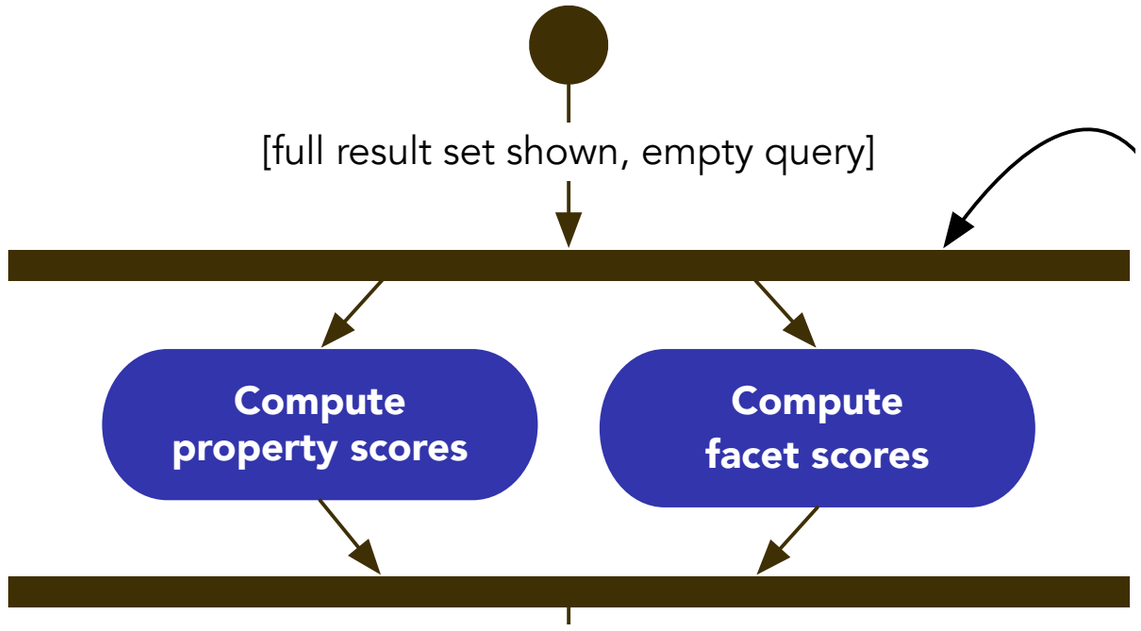
- Selected facets is the query
- Disjunctive vs conjunctive semantics, e.g.
  - `Brand:Apple`
  - `Brand:Samsung`
  - `Color:Black`

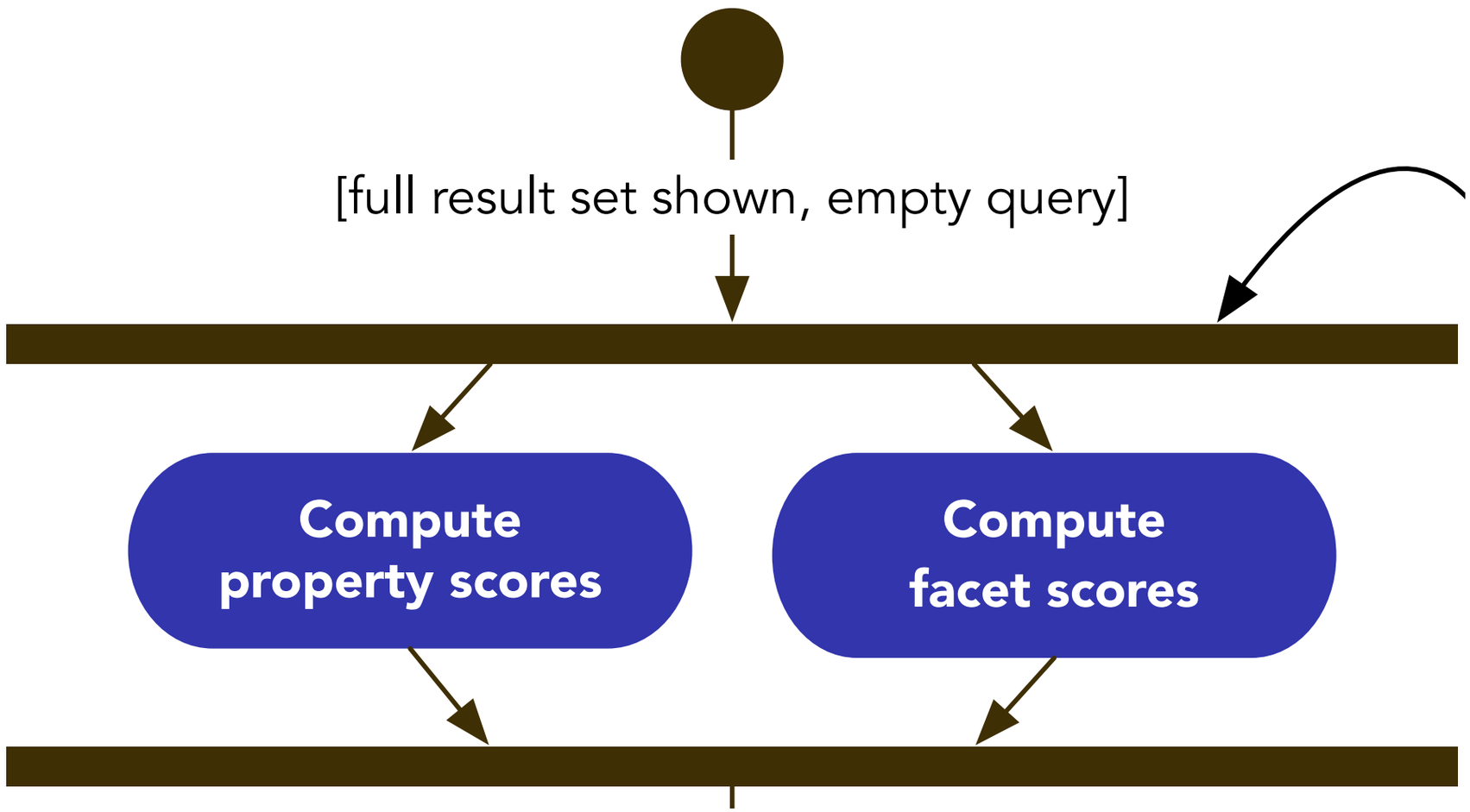
Approach

Search sessions

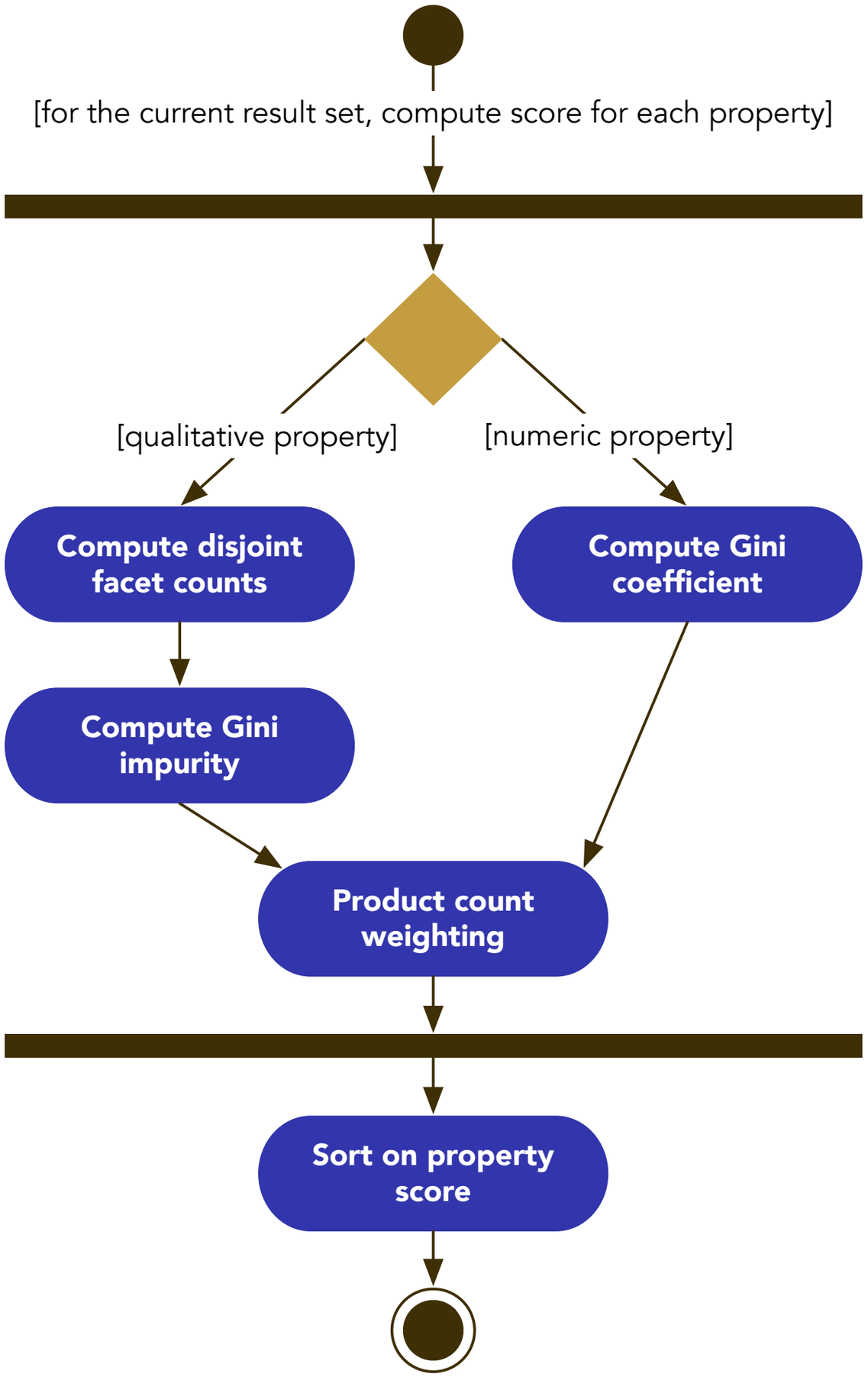
- User can perform drill-down or drill-up
- Roll-up can occur when a user:
  - deselect last remaining facet
  - selects an additional qualitative facet
  - broadening numeric range


[full result set shown, empty query]
Compute property scores
Compute facet scores
Present ordered properties and facets
[result size too large user performs drill-down]
Update result set using query
[result set small enough, user scans products]
[product not found, user performs roll-up]
[user finds target product]


[full result set shown, empty query]
Compute property scores
Compute facet scores


[full result set shown, empty query]
Compute property scores
Compute facet scores

Compute property scores



Approach

Property Scores (qualitative properties)

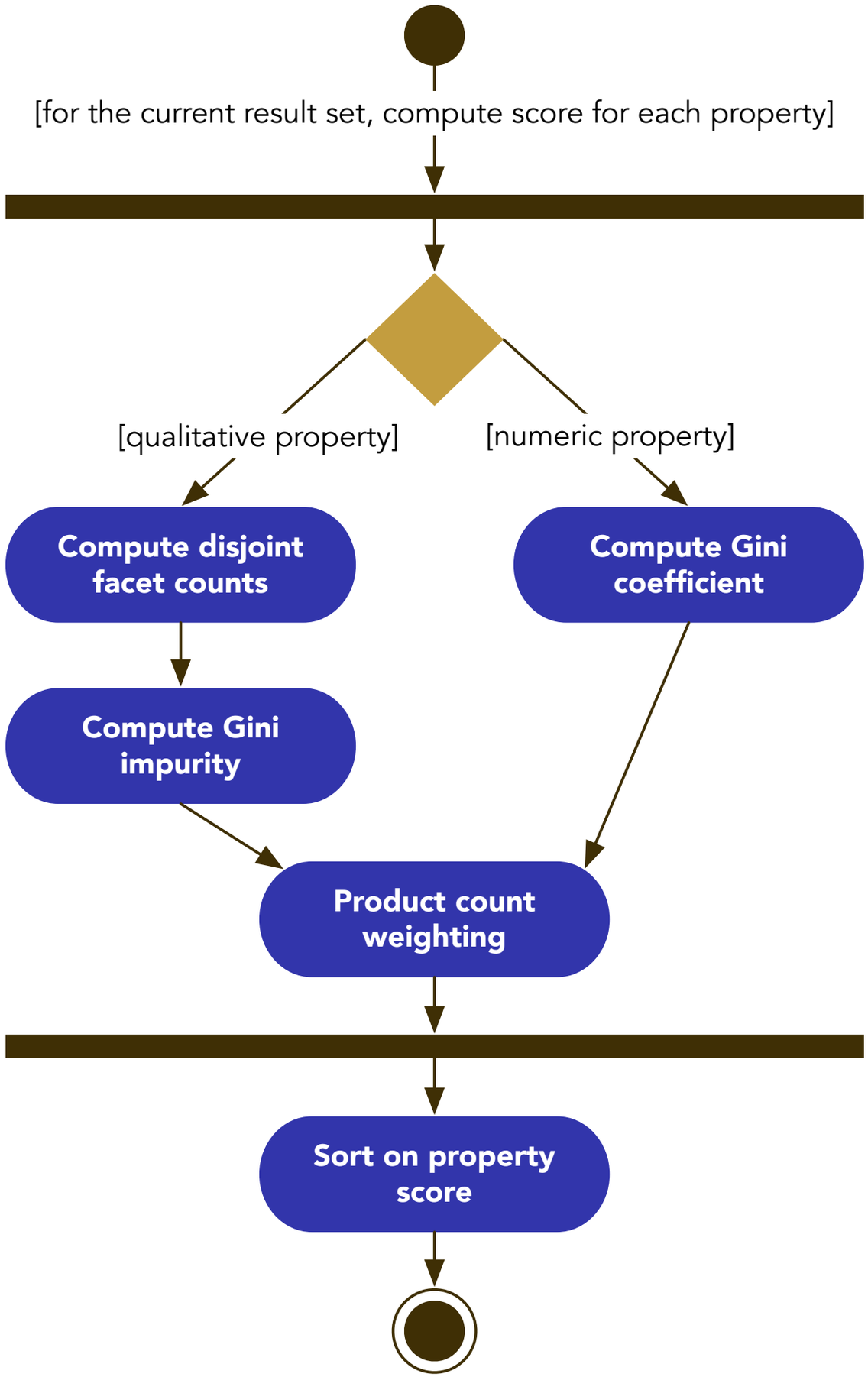
$$\text{count}(f, q) = |D_q \cap D_f| = \sum_{d \in D_q} \begin{cases} 1 & \text{if } f \in F_d \\ 0 & \text{if } f \notin F_d \end{cases} \tag{1}$$

$$\text{disjointCount}(f, q) = \sum_{d \in D_q} \begin{cases} 1 & \text{if } F_p \cap F_d \equiv \{f\} \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

Approach

Property Scores (qualitative properties)

$$\text{giniImpurity}(p, q) = 1 - \sum_{f \in F_p} \left(\frac{\text{disjointCount}(f, q)}{\sum_{g \in F_p} \text{disjointCount}(g, q)} \right)^2 \quad (3)$$


[for the current result set, compute score for each property]
[qualitative property]
[numeric property]
Compute disjoint facet counts
Compute Gini coefficient
Compute Gini impurity
Product count weighting
Sort on property score

Approach

Property Scores (numeric properties)

- use distribution of values
- prefer properties with widely dispersed values
- we employ the Gini coefficient

Approach

$$\text{giniCoefficient}(p, q) = \frac{1}{m}\left(m + 1 - 2\left(\frac{\sum_{i=1}^{m}(m+1-i)f_i}{\sum_{i=1}^{m} f_i}\right)\right) \quad (4)$$

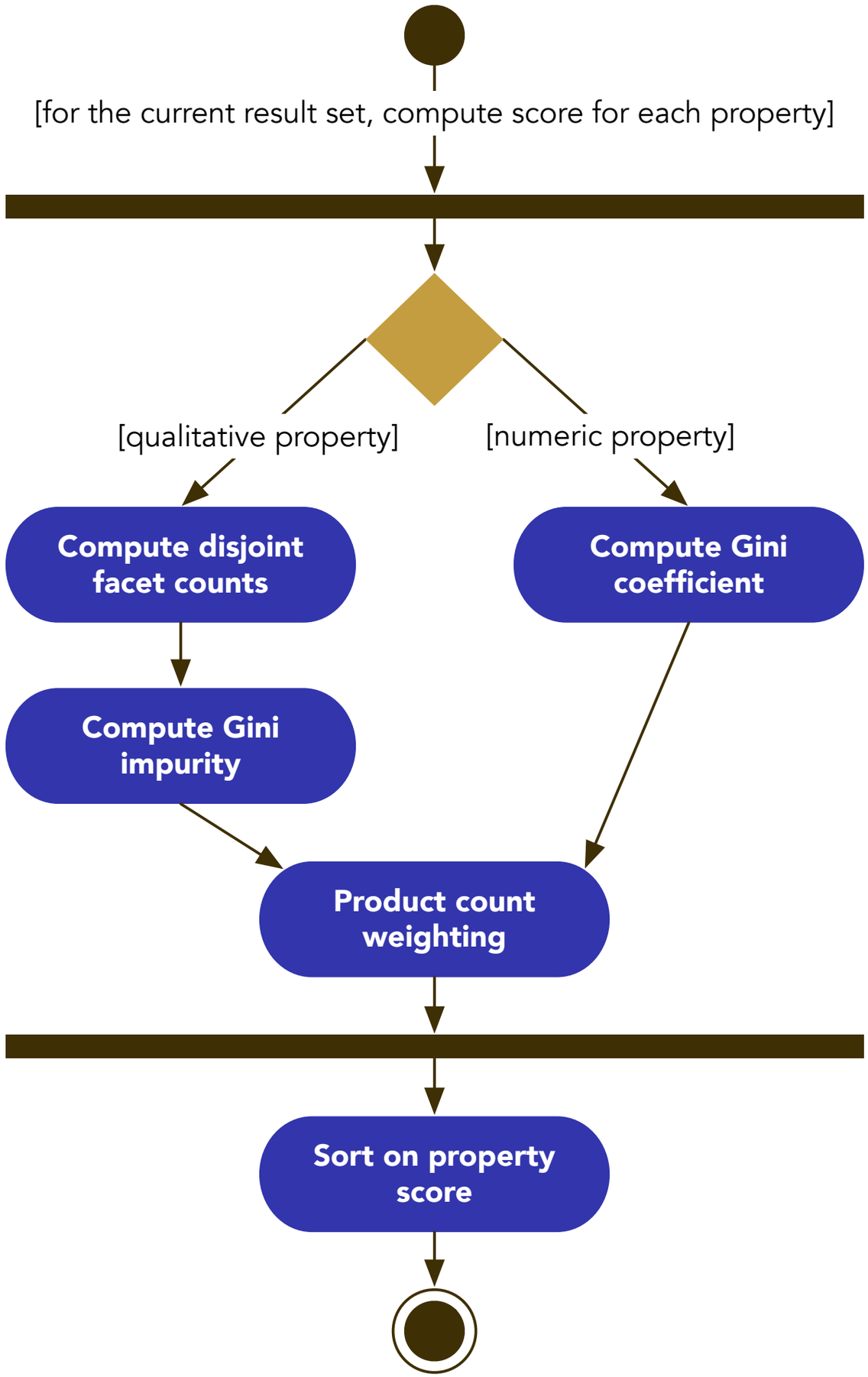
$$= \frac{2\sum_{i=1}^{m} if_i}{m\sum_{i=1}^{m} f_i} - \frac{m+1}{m}$$

$$\text{given } f_i \in F_p^* \text{ for } i = 1 \text{ to } m$$

$$F_p^* = \{f_i \mid f_i \in F_p \cap F_d,\ d \in D_q,\ f_i \leq f_{i+1}\}$$

$$m = |F_p^*|$$

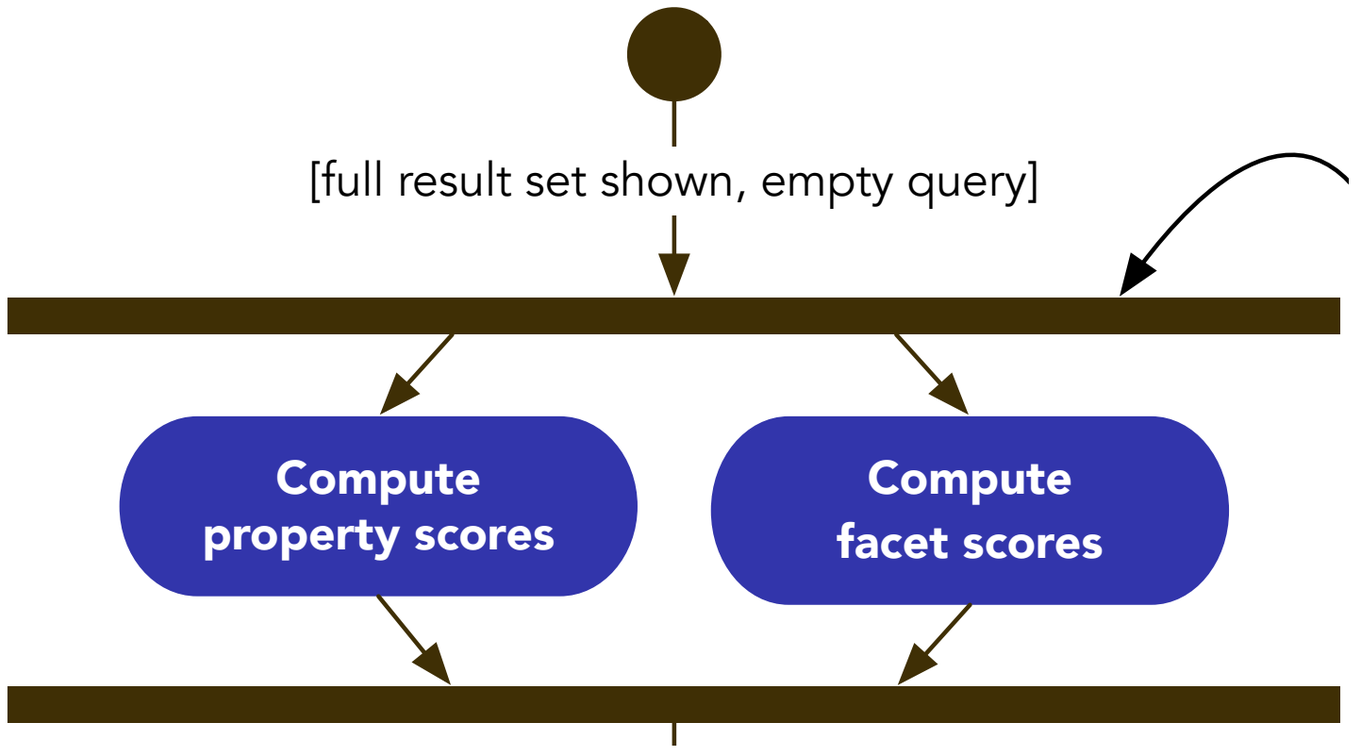
$$p \in P_{\text{quantitative}}$$

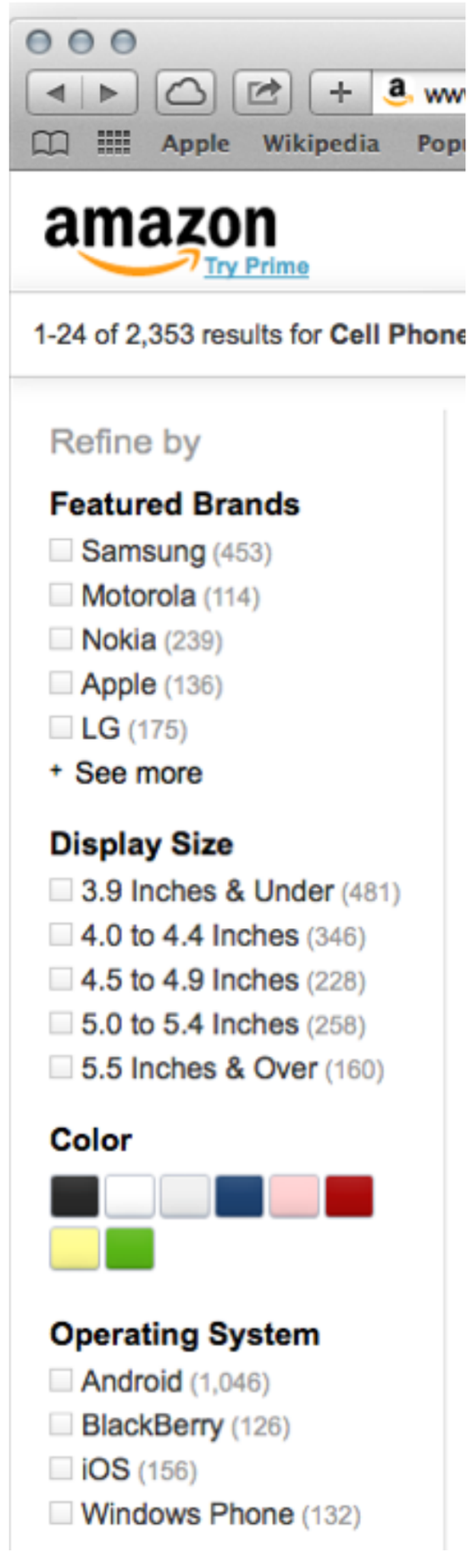

[for the current result set, compute score for each property]
[qualitative property]
[numeric property]
Compute disjoint facet counts
Compute Gini coefficient
Compute Gini impurity
Product count weighting
Sort on property score

Approach

Product count weighting

$$\text{propertyScore}(p, q) = gini(p, q) \cdot \sum_{f \in F_p} \frac{\text{disjointCount}(f, q)}{|D_q|} \tag{5}$$

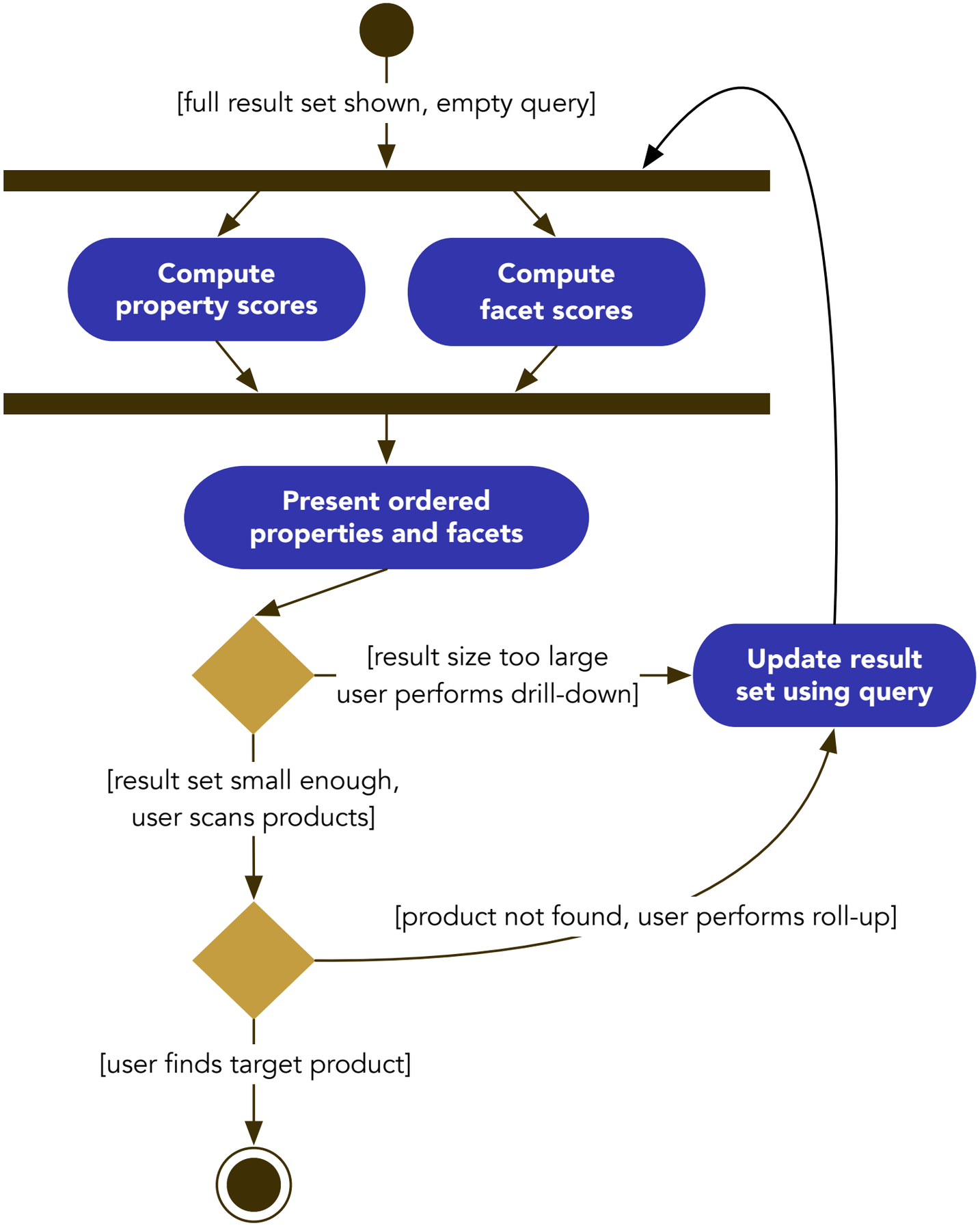

[full result set shown, empty query]
Compute property scores
Compute facet scores


Apple
Wikipedia
amazon
Try Prime
1-24 of 2,353 results for Cell Phone
Refine by
Featured Brands
Samsung (453)
Motorola (114)
Nokia (239)
Apple (136)
LG (175)
+ See more
Display Size
3.9 Inches & Under (481)
4.0 to 4.4 Inches (346)
4.5 to 4.9 Inches (228)
5.0 to 5.4 Inches (258)
5.5 Inches & Over (160)
Color
Operating System
Android (1,046)
BlackBerry (126)
iOS (156)
Windows Phone (132)

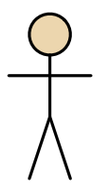
Approach

Facet scores computation

- For numeric properties, we ignore facet scores
    - usually represented with sliders in UI's
- For qualitative properties, rank descending on the facet count
    - increases chance that a facet matching the target product will be selected


[full result set shown, empty query]
Compute property scores
Compute facet scores
Present ordered properties and facets
[result size too large user performs drill-down]
Update result set using query
[result set small enough, user scans products]
[product not found, user performs roll-up]
[user finds target product]

Evaluation



Drill-Down Model

Modeling user behavior in the simulations:

- Least Scanning
- Best Facet
- Combined

+

Ordering Scheme

Representing the system's approach for ranking facets:

- Expert-Based
- Greedy Count
- Kim et al.
- Our Approach

+

Target Product

The product to find for each experiment:

- Apple iPhone 4 16GB
- LG Nexus 4 8GB
- …

(794 products)

Repetitions

To reduce stochastic effects each experiment is repeated:

50x

Performance Measures

User Effort:

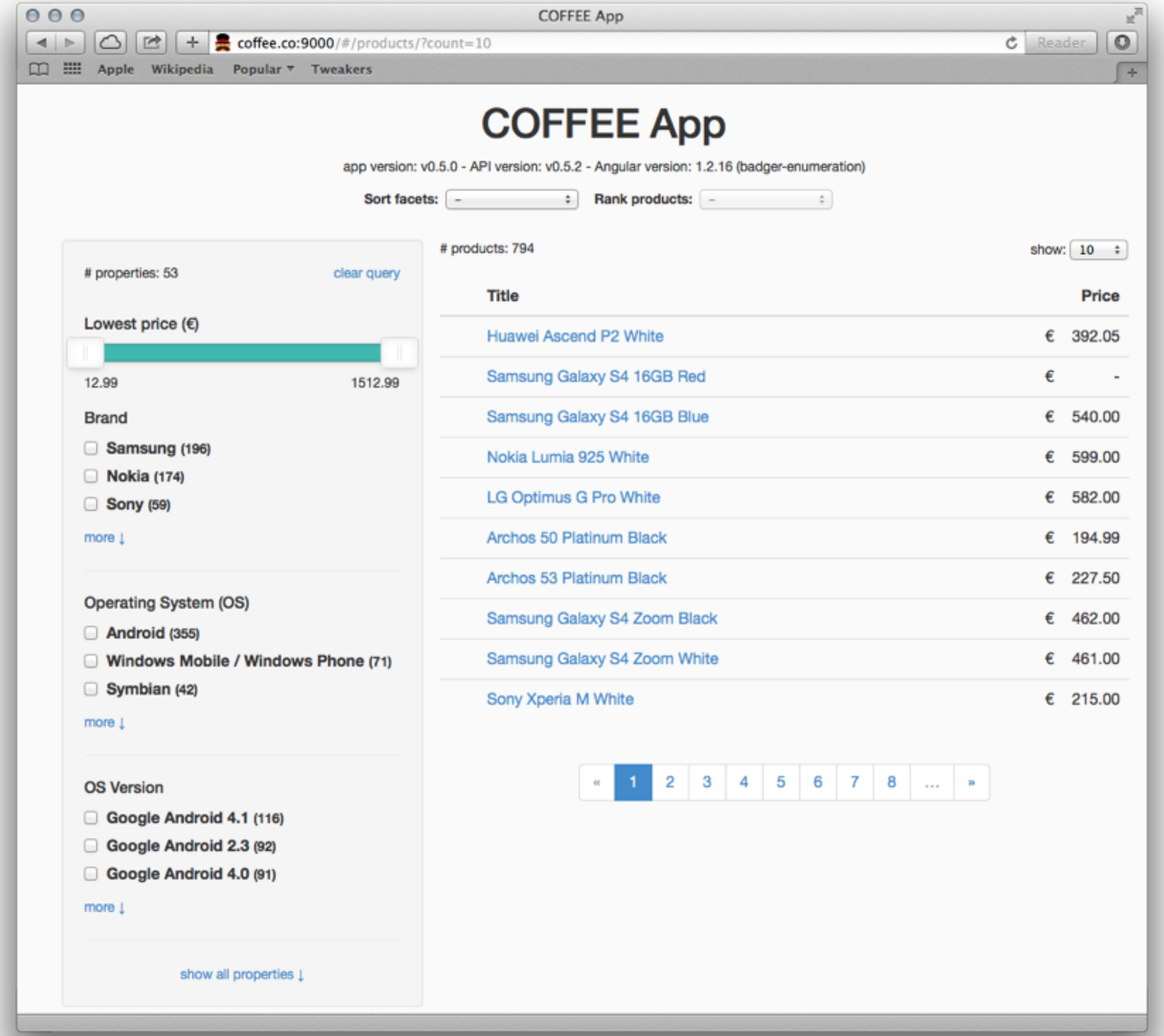
- Click Effort
- Property Scan Effort
- Value Scan Effort

Other measures:

- Computation Time
- # Roll-Ups
- % Successful Sessions

Evaluation

- Tweakers PriceWatch data set
- 794 mobile phones
- 53 properties and 1,816 facets
  - 348 qualitative
  - 1,468 numeric
- Over 150,000 experiments run on a cluster
- Implemented demo application


COFFEE App
coffee.co:9000/#/products/?count=10
Reader
Apple
Wikipedia
Popular
Tweakers
COFFEE App
app version: v0.5.0 - API version: v0.5.2 - Angular version: 1.2.16 (badger-enumeration)
Sort facets:
Rank products:
properties: 53
clear query
Lowest price (€)
12.99
1512.99
Brand
Samsung (196)
Nokia (174)
Sony (59)
more ↓
Operating System (OS)
Android (355)
Windows Mobile / Windows Phone (71)
Symbian (42)
more ↓
OS Version
Google Android 4.1 (116)
Google Android 2.3 (92)
Google Android 4.0 (91)
more ↓
show all properties ↓
products: 794
show: 10
Title
Price
Huawei Ascend P2 White € 392.05
Samsung Galaxy S4 16GB Red € -
Samsung Galaxy S4 16GB Blue € 540.00
Nokia Lumia 925 White € 599.00
LG Optimus G Pro White € 582.00
Archos 50 Platinum Black € 194.99
Archos 53 Platinum Black € 227.50
Samsung Galaxy S4 Zoom Black € 462.00
Samsung Galaxy S4 Zoom White € 461.00
Sony Xperia M White € 215.00
« 1 2 3 4 5 6 7 8 ... »

	Ordering Scheme			
	Expert-Based	Greedy Count	Kim et al.	Our approach
user effort:				
# clicks (X_c)	4.0	28.2	19.7	2.3
# clicks std dev	1.24	18.65	14.04	0.68
prop scan effort (X_p)	0.0538	0.1914	0.0630	0.0267
prop scan effort std dev	0.0273	0.0891	0.0351	0.0124
facet scan effort (X_f)	0.1462	0.2438	0.4550	0.2111
facet scan effort std dev	0.0908	0.0952	0.1516	0.1718
other measures:				
computation time (ms)	4	23,386	49,818	187
computation time std dev	3.7	26,832.4	45,129.9	74.9
successful sessions (%)	100.00%	100.00%	100.00%	100.00%

(a) Least Scanning Drill-Down Model

	Ordering Scheme			
	Expert-Based	Greedy Count	Kim et al.	Our approach
user effort:				
# clicks (X_c)	1.5	1.5	1.5	1.5
# clicks std dev	0.52	0.52	0.52	0.52
prop scan effort (X_p)	0.3474	0.7232	0.5804	0.2399
prop scan effort std dev	0.2607	0.2091	0.1939	0.2257
facet scan effort (X_f)	0.4659	0.4796	0.4946	0.4547
facet scan effort std dev	0.2730	0.2736	0.2695	0.2764
other measures:				
computation time (ms)	2	25	1,507	160
computation time std dev	0.9	213.2	638.1	61.9
successful sessions (%)	100.00%	100.00%	100.00%	100.00%

(b) Best Facet Drill-Down Model

	Ordering Scheme			
	Expert-Based	Greedy Count	Kim et al.	Our approach
user effort:				
# clicks (X_c)	30.7	62.9	59.8	18.8
# clicks std dev	20.05	27.98	20.01	9.77
prop scan effort (X_p)	0.1220	0.1681	0.1524	0.2268
prop scan effort std dev	0.0232	0.0255	0.0297	0.0261
facet scan effort (X_f)	0.3904	0.4842	0.5443	0.3075
facet scan effort std dev	0.0599	0.1100	0.0325	0.0308
other measures:				
computation time (ms)	16	118,155	113,336	2,843
computation time std dev	12.6	72,772.1	53,871.0	2,094.0
# rollups mean	10.7	10.0	16.6	6.2
successful sessions (%)	90.96%	64.00%	79.53%	99.07%

(c) Combined Drill-Down Model

	Ordering Scheme			
	Expert-Based	Greedy Count	Kim et al.	Our approach
user effort:				
# clicks (X_c)	4.0	28.2	19.7	2.3
# clicks std dev	1.24	18.65	14.04	0.68
prop scan effort (X_p)	0.0538	0.1914	0.0630	0.0267
prop scan effort std dev	0.0273	0.0891	0.0351	0.0124
facet scan effort (X_f)	0.1462	0.2438	0.4550	0.2111
facet scan effort std dev	0.0908	0.0952	0.1516	0.1718
other measures:				
computation time (ms)	4	23,386	49,818	187
computation time std dev	3.7	26,832.4	45,129.9	74.9
successful sessions (%)	100.00%	100.00%	100.00%	100.00%

(a) Least Scanning Drill-Down Model

	Ordering Scheme			
	Expert-Based	Greedy Count	Kim et al.	Our approach
user effort:				
# clicks (X_c)	1.5	1.5	1.5	1.5
# clicks std dev	0.52	0.52	0.52	0.52
prop scan effort (X_p)	0.3474	0.7232	0.5804	0.2399
prop scan effort std dev	0.2607	0.2091	0.1939	0.2257
facet scan effort (X_f)	0.4659	0.4796	0.4946	0.4547
facet scan effort std dev	0.2730	0.2736	0.2695	0.2764
other measures:				
computation time (ms)	2	25	1,507	160
computation time std dev	0.9	213.2	638.1	61.9
successful sessions (%)	100.00%	100.00%	100.00%	100.00%

(b) Best Facet Drill-Down Model

	Ordering Scheme			
	Expert-Based	Greedy Count	Kim et al.	Our approach
user effort:				
# clicks (X_c)	30.7	62.9	59.8	18.8
# clicks std dev	20.05	27.98	20.01	9.77
prop scan effort (X_p)	0.1220	0.1681	0.1524	0.2268
prop scan effort std dev	0.0232	0.0255	0.0297	0.0261
facet scan effort (X_f)	0.3904	0.4842	0.5443	0.3075
facet scan effort std dev	0.0599	0.1100	0.0325	0.0308
other measures:				
computation time (ms)	16	118,155	113,336	2,843
computation time std dev	12.6	72,772.1	53,871.0	2,094.0
# rollups mean	10.7	10.0	16.6	6.2
successful sessions (%)	90.96%	64.00%	79.53%	99.07%

(c) Combined Drill-Down Model

Conclusions

- We proposed an facet ordering approach
- Two Gini-based measures for qualitative and numeric properties
- Compared to other (automatic) approaches:
  - faster
  - needs less roll-ups
  - higher % successful sessions