

# Personalization and User Modeling

Overlay User Modeling, Entropy and Overlap of User Profiles

Patrick Siehndel

IVS Semantic Web Group

19.05.2014

# Exercise 1: Overlay User Modeling

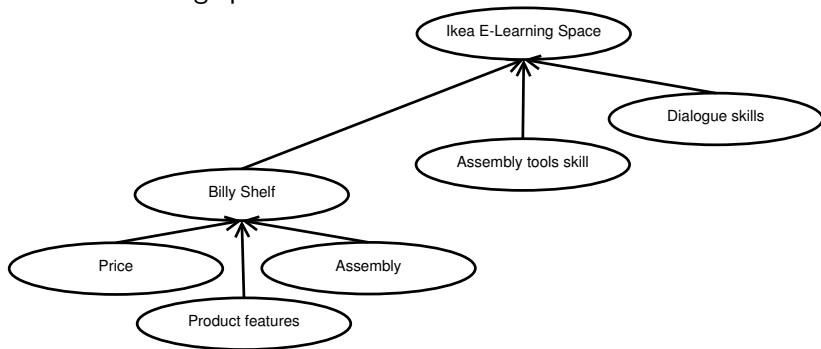
Ikea plans to offer online training courses for their employees. They want to train their in-house sales stuff, telephone operator and assemblers.

- Please create a domain model that describes the Ikea domain knowledge: Therefore, define the concepts and prerequisites of the Ikea learning space.
- Please select two of the employees and specify a corresponding overlay model.
- Which weighting schemes can be used to specify the degree of knowledge?

# Solution 1: Overlay User Modeling

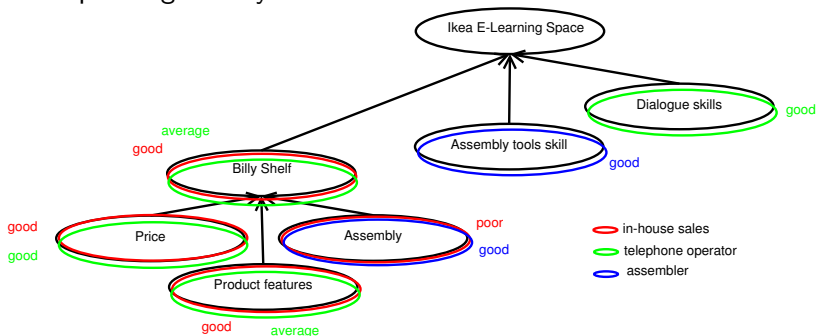
Ikea plans to offer online training courses for their employees. They want to train their in-house sales stuff, telephone operator and assemblers.

- Please create a domain model that describes the Ikea domain knowledge: Therefore, define the concepts and prerequisites of the Ikea learning space.



# Solution 1: Overlay User Modeling

- Please select two of the employees and specify a corresponding overlay model.



- Which weighting schemes can be used to specify the degree of knowledge?

See lecture slide 8: The binary overlay model does not fit with the defined domain model for the setting – for example, both assemblers and in-house sales stuff should have an understanding about how to assemble the furniture. However, for sales stuff basic skills are sufficient while assemblers need a detailed knowledge about the assembly. Uncertainty-based and numeric weighting scheme might be a bit too fine-grained so that we preferably choose qualitative schema.

## Exercise 2 : Entropy and Overlap

Given are the following user profiles:

profile	tag (frequency)
flickr-bob	hannover (8) italy (4)
stumble-bob	research (8) semantic web (8)
delicious-bob	semantic web (10) social web (5) hannover (3) user modeling (3)
mypes-bob (aggregated)	semantic web (14) hannover (11) italy (8) research (8) social web (5) user modeling (3)

- 1 Calculate the entropy for the given user profiles.
- 2 Calculate the pairwise overlap between the first three given user profiles.
- 3 Give an example of two user profiles where the entropy of one of the single user profiles is higher than the entropy of the combined user profile.

# Solution 2: Entropy and Overlap

- 1 Calculate the entropy for the given user profiles.

To calculate the entropy we need to determine the occurrence probability of each keyword of a keyword-based profile. This can be calculated by

$p(x) = \frac{\text{freq}(x)}{|\sum_{t \in T} \text{freq}(t)|}$ . For the occurrence probability of hannover in the flickr profile we get  $p(\text{hannover}) = \frac{\text{freq}(\text{hannover})}{\text{freq}(\text{hannover}) + \text{freq}(\text{italy})} = \frac{8}{8+4} \approx 0.67$ .

The entropy can be calculated as

$$\begin{aligned} \text{entropy}(\text{flickr} - \text{bob}) &= \sum_{t \in \{\text{hannover}, \text{italy}\}} p(t) * -\log_2(p(t)) \\ &= p(\text{hannover}) * -\log_2(p(\text{hannover})) + p(\text{italy}) * -\log_2(p(\text{italy})) \\ &= 0.67 * 0.59 + 0.33 * 1.59 \\ &= 0.92 \end{aligned}$$

$$\begin{aligned} \text{entropy}(\text{stumble} - \text{bob}) &= 1 \\ \text{entropy}(\text{delicious} - \text{bob}) &= 1.8 \\ \text{entropy}(\text{mypes} - \text{bob}) &= 2.44 \end{aligned}$$

## Solution 2: Entropy and Overlap

- 2 Calculate the pairwise overlap between the first three given user profiles. To calculate the overlap we first need to count the number of tags that occur in both user profiles. E.g.  $T_{flickr-bob} \cap T_{stumble-bob} = 0$ ,  $T_{flickr-bob} \cap T_{delicious-bob} = 1$ . The overlap is then calculated as follows:

$$\begin{aligned} \text{overlap}(flickr - bob, stumble - bob) &= \\ \frac{1}{2} * \left( \frac{T_{flickr-bob} \cap T_{stumble-bob}}{T_{flickr-bob}} + \frac{T_{flickr-bob} \cap T_{stumble-bob}}{T_{stumble-bob}} \right) &= \\ \frac{1}{2} * \left( \frac{0}{2} + \frac{0}{2} \right) &= 0 \\ \\ \text{overlap}(flickr - bob, delicious - bob) &= \frac{3}{8} \\ \text{overlap}(stumble - bob, delicious - bob) &= \frac{3}{8} \end{aligned}$$

- 3 Give an example of two user profiles where the entropy of one single user profile is higher than the entropy of the combined user profile.

We use as input the stumble-bob user profile. To lower the entropy we need to make the profile more predictable. If we add as second user profile a profile that contains the word *research* eighth times we end up with the same word distribution as in flickr-bob and therefore receive the same lower entropy.