

Analysis of Subjective Conceptualizations towards Collective Conceptual Modelling

Fumiko Kano Glückstad*¹ Tue Herlau*² Mikkel N. Schmidt*² Morten Mørup*²

*¹ Copenhagen Business School, Dept. of International Business Communication

*² Technical University of Denmark, DTU Compute, Section for Cognitive Systems

This work is conducted as a preliminary study for a project where individuals' conceptualizations of domain knowledge will thoroughly be analyzed across 150 subjects from 6 countries. The project aims at investigating how humans' conceptualizations differ according to different types of mother languages, cultural backgrounds, gender, generations, etc., when domain-specific terms are expressed in a common language, i.e. English. In this work, we analyze a publicly available dataset [De Deyne, 2008] representing semantic structures of domain knowledge possessed by four subjects. The application of a non-parametric relational model, Infinite Relational Model [Kemp, 2006] co-clusters concept-feature relations, which identifies a common semantic structural grid across the four subjects considered. Through this common grid, the individual semantic structures possessed by the respective subjects are further contrasted to each other. The results show that the identified common grid visualizes subject specific patterns among the extracted concept-feature clusters. The work further discusses our future perspectives for modeling a collective ontology across subjects according to different subjects' profiles.

1. Introduction

In the history of computer sciences and the philosophy of ontology, knowledge has often been considered as static at a certain temporal point. In this mindset, a standardized ontology defined by a third party [Gruber 92][Madsen 04][Declerck 10] has been playing an important role to establish a common ground that can facilitate mutual understandings among different communicating parties. However, the patterns of communication in our real life cannot be accounted for solely by standardized knowledge that is defined in a top-down manner. Scientists in semiotics [Pierce 58] [Durst-Andersen 11], cognitive pragmatics [Sperber 86] and cognitive sciences [Rosch 75][Storms 97] put emphasis on mental representations of knowledge that are constructed differently due to differences in our close environment and our different cognitive abilities. Hence, it is challenging but also fascinating to explore how differences in subjective conceptualization should be treated for the purpose of cross-cultural communication and collaboration. This work attempts to identify a common semantic structural grid where different patterns of conceptualizations are represented in a bottom-up manner.

In this work, an unsupervised learning method known as Infinite Relational Model (IRM) [Kemp 06] is employed. A framework of applying the IRM in this work is primary based on the design described in [Mørup 10]. The IRM is used to co-cluster concept-feature relations, which identifies a common semantic structural grid across multiple human subjects. This means that the identified concept- and feature clusters that are partitioned according to the

common grid are consistent across subjects, while the interactions between concepts and features are subject specific. Through this optimally identified grid, differences across individual semantic structures possessed by the respective subjects can be inspected. For visualization we apply an extended version of IRM, the normal Infinite Relational model (n-IRM) [Herlau 12] in order to further organize the subject specific patterns of interactions obtained from the optimally identified common grid. This enables one to analyze different patterns of conceptualization by the four subjects.

In the next section, the methods employed in this work are reviewed. Section 3 presents the results obtained by the IRM and the n-IRM, followed by the data analysis in Section 4. Section 5 further discusses future research perspectives, followed by the conclusions in Section 6.

2. Methods

2.1 Semantic Structures

In cognitive sciences, a semantic structure of domain knowledge possessed by each human subject is often represented by a binary matrix consisting of objects (concepts) and attributes (features). De Deyne et al. [De Deyne 08] developed and verified a data matrix frame representing the domain of *clothing*. In [De Deyne 08], the reliability of concepts to be listed in the matrix has been validated with a number of human subjects based on multiple approaches, among others: typicality ratings, goodness ratings, exemplar generation frequency, exemplar and category associative strength, familiarity ratings, imaginability ratings, and pairwise similarity ratings. After the validation of the concepts list, De Deyne et al. [De Deyne 08] implemented feature-generation tasks and feature-importance ratings for each concept. The generated features for each concept are listed in a concept-feature matrix and its applicability has eventually been assessed by four individual hu-

man subjects in [De Deyne 08]. The clothing matrix framework developed in this way by [De Deyne 08] consists of 29 concepts, e.g. *blouse*, *pants*, *dress*, and 258 features, e.g. *protects against the cold*, *made of silk*, *gives a business like impression*. For every feature-concept pair, the four human subjects respectively assigned a binary value 1, if a feature is applicable to the concept in question, otherwise 0 was assigned. In this way, four individual semantic structures, i.e. the concept-feature matrices, assessed by the respective subjects have been prepared in [De Deyne 08].

2.2 The Infinite Relational Model

The goal of the IRM described by [Kemp 06] is to *partition each type into clusters, where a good set of partitions allows relationships between entities to be predicted by their cluster assignments*. A key feature of the IRM is to *automatically choose an appropriate number of clusters using a prior that favors small numbers of clusters, but has access to a countably infinite collection of clusters* [Kemp 06]. This work employs the simplest model: dealing with two types T_1 and T_2 and a two-place relation $R : T_1 \times T_2 \rightarrow \{0, 1\}$, where T_1 and T_2 respectively corresponds to the concepts and their features in the clothing domain. Since we simultaneously analyze multiple matrices, we employed the multi-array IRM framework that is similar to the one described in [Mørup 10]. Thus, the IRM is applied to the multi-array binary matrices representing semantic structures of the four human subjects prepared in [De Deyne 08].

The principle of generating clusters in the IRM [Kemp 06] is based on a distribution over partitions induced by a so-called Chinese Restaurant Process (CRP) [Pitman 02]. The CRP induce a partition by first creating a single cluster with only one object, and then sequentially adding the other objects to the partition by letting the i^{th} object either select a new cluster with probability $\frac{\gamma}{i-1+\gamma}$, or an existing cluster a with probability $\frac{m_a}{i-1+\gamma}$. Here, m_a is the number of objects already assigned to cluster a , and $\gamma > 0$ is a parameter.

The clusterings of the two types T_1 and T_2 are respectively defined as $\mathbf{z}^{(1)}$ and $\mathbf{z}^{(2)}$. For the s^{th} binary matrix having a link is stated as $p(R_{ij}^{(s)} = 1 | \mathbf{z}^{(1)}, \mathbf{z}^{(2)}) = \eta_{ab}^{(s)}$, where $a = z_i^{(1)}$ and $b = z_j^{(2)}$ are the clusters i and j belong to. Accordingly the generative model is defined as:

$$\begin{aligned} \mathbf{z}^{(1)} &\sim \text{CRP}(\gamma^{(1)}) && \textit{first mode}, \\ \mathbf{z}^{(2)} &\sim \text{CRP}(\gamma^{(2)}) && \textit{second mode}, \\ \eta_{ab}^{(s)} &\sim \text{Beta}(\beta_0^+, \beta_0^-) && \textit{interactions}, \\ R_{ij}^{(s)} &\sim \text{Bernoulli} \left(\eta_{z_i^{(1)} z_j^{(2)}}^{(s)} \right) && \textit{links}. \end{aligned}$$

In this generative model, the first- and the second lines respectively partition concepts and features into $\mathbf{z}^{(1)}$ and $\mathbf{z}^{(2)}$ clusters. In the third line, how concept clusters relate to feature clusters are defined for each human subject s^{th} by $\eta_{ab}^{(s)}$. Finally, $R_{ij}^{(s)}$ draws binary relations between concepts and their features based on the extracted clusters and their subject specific interactions.

In the above model, we set prior parameters $\beta_0^\pm = 1$, and $\gamma^{(\alpha)} = \log(J_\alpha)$ where J_α is the number of concepts in each

mode. Inference in the model is performed by Gibbs sampling along with a split-merge Metropolis Hastings framework as described in [Mørup 10]. The number of iteration is chosen as 1000 where the first 500 are used for burnin. The solutions displayed in the followings are based on the single realization with highest likelihood.

2.3 The normal Infinite Relational Model

The n-IRM [Herlau 12] is a generalization of the IRM to continuous data. The name *normal* IRM implies that the model employs normal distribution as likelihood function and normal-inverse gamma priors instead of the Bernoulli likelihood and beta priors. Otherwise the process of generating clusters from the CRP and letting the value of the relation (now allowed to be any real number) be conditionally independent given the clustering is the same. The n-IRM employs a Bayesian generative model for dividing each mode of the s^{th} relational matrix $W_{ij}^{(s)}$ into clusters with a corresponding interaction matrix between the groups of size $n^{(1)} \times n^{(2)}$. The elements of each $n_a \times n_b$ submatrix is parameterized by a mean intensity $m_{ab}^{(s)}$ and precision $\lambda_{ab}^{(s)} = \sigma_{ab}^{-2}$, which are given a Normal-Gamma prior. Accordingly, the generative model is defined as

$$\begin{aligned} \mathbf{z}^{(1)} &\sim \text{CRP}(\gamma^{(1)}) && \textit{first mode}, \\ \mathbf{z}^{(2)} &\sim \text{CRP}(\gamma^{(2)}) && \textit{second mode}, \\ \lambda_{ab}^{(s)} &\sim \text{Gam}(\alpha_0, \text{rate} = \beta_0) && \textit{precision}, \\ m_{ab}^{(s)} &\sim \text{Normal}(m_0, (\kappa_0 \lambda_{ab}^{(s)})^{-1}) && \textit{mean}, \\ R_{ij}^{(s)} &\sim \text{Normal}(m_{z_i^{(1)} z_j^{(2)}}^{(s)}, \lambda_{z_i^{(1)} z_j^{(2)}}^{(s)}) && \textit{links}. \end{aligned}$$

The operation for the generative process is further described in [Herlau 12].

In this work, the n-IRM is applied to the multi-array eta-value $\eta_{(ab)}$ matrices, where $\eta_{(ab)}$ is the highest likelihood of relations between $\mathbf{z}^{(1)}$ and $\mathbf{z}^{(2)}$. The prior values for the n-IRM was set to $\kappa_0 = 1$, $\alpha_0 = 15$ and $\beta_0^{-1} = 1$ to reflect the scale of the η values. The iteration defined here is the same as the IRM implementation, i.e. 1000 where the first 500 iterations are used as burnin for the sampler.

3. Results

Figure 1 depicts the original binary matrices representing the semantic structures of the four subjects prepared by [De Deyne 08]. The y axis, x axis and blue dots of each graph respectively represent 29 clothing concepts, 258 characteristic features and their binary values where each subject assigned 1 if a feature is applicable to a concept in question. These four graphs indicate that the semantic structures possessed by the four subjects are slightly different. Figure 2 illustrates individual semantic structures of the four subjects which are contrasted via a common grid optimally computed by the IRM. The graphs show that the 29 clothing concepts and the 258 features are respectively partitioned into 6 and 36 clusters. In Figure 2, the lines partitioning the concepts and features are considered as an optimal grid that is common across the four subjects.

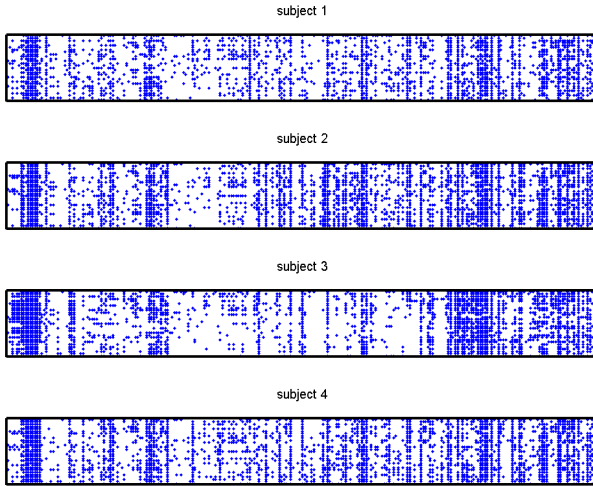


Figure 1: Original binary matrices representing the semantic structures of the four subjects

When observing individual blocks of the grid, different patterns across the four subjects become visible. The four plots in the left in Figure 3 shows the density (η -values) of each interaction between a concept-cluster and a feature-cluster, where the darkest gray indicates the highest density and vice versa. These four η matrices are further partitioned by the n-IRM in the plots of the right column. Here, the n-IRM has been applied in order to co-cluster the η -value matrices displayed in the left column for further visualizing what group of feature clusters influence the formation of the concept-clusters as well as the differentiation of the semantic structures possessed by the four subjects.

Figure 4 illustrates the results of the η -value partitions achieved by the n-IRM across the four subjects. For example, the four plots of the mean values of the normal distributions indicate that *feature-cluster cluster 3* (referred to as *fc3*) consisting of feature clusters f5, f9, f13, f21, f26 and f27, and *fc5* consisting of feature clusters f1, f2, f3, and f4 are relatively uniform across the four subjects. On the other hand, *fc2* (feature clusters: 12, 19, 22, 31, 32, 33, 34) and *fc4* (feature clusters: 16, 23, 25, 30, 35, 36), *fc6* (feature cluster: 28), *fc7* (feature cluster: 7), *fc8* (feature cluster: 29) are irregular across the four subjects. The four plots in the right giving the variance of the normal distributions in 4 show that, for example *fc1* (feature clusters: 6, 8, 10, 11, 14, 15, 17, 18, 20, 24), *fc2* (feature clusters: 12, 19, 22, 31, 32, 33, 34) and *fc4* (feature clusters: 16, 23, 25, 30, 35, 36) result in larger variances for all the four subjects. This indicates that these *feature-cluster clusters* consist of the feature clusters that have stronger influence on distinguishing concept clusters in each semantic structure.

In this way, Figures 2, 3 and 4 overview the patterns of the semantic structures possessed by the four subjects by contrasting an interaction between a concept-cluster and a feature-cluster through the common grid that have optimally been computed by the IRM. In the next section,

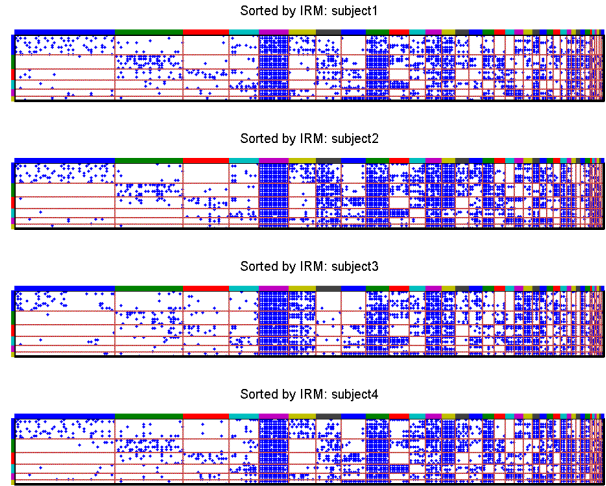


Figure 2: Individual semantic structures contrasted via a common grid

the contents of the obtained clusters are further analyzed in details.

4. Analysis

Figures 5, 6 and 7 summarize the aforementioned results with concrete concept- and feature labels. The concept-clusters obtained across the four subjects from the IRM analysis are: cluster 1 (hat, boots, beanie, cap, belt, shoes, scarf, socks, mittens); cluster 2 (bra, dress, skirt, panties, top, bathing suit); cluster 3 (tracksuit, pyjamas, dungarees, shorts, t-shirt); cluster 4 (blouse, tie, shirt, suit); cluster 5 (coat, sweater, pullover); and cluster 6 (pants, jeans). From the common sense intuition, the IRM has reasonably partitioned 29 concepts into 6 clusters that could respectively characterized as: 1) accessories; 2) items targeted for female, 3) casual items, 4) formal items, 5) items used outside, and 6) items covering legs. This motivate us to further investigate how these clusters have been formed based on the distribution of 258 features whose associations are individually assessed by the four subjects.

In Figures 5, 6 and 7, the IRM results, i.e. 6 concept-clusters and 36 feature-clusters, are re-ordered computed by the n-IRM.

Figure 5 depicts *fc1* that consists of feature clusters f6, f8, f10, f11, f14, f15, f17, f18, f20 and f24. As described in the previous section, the variances in *fc1* are high, while the mean values are low in all the four subjects. This indicates that these 10 feature-clusters have, in general, strong influences on the formation of specific concept-clusters. For example, f18 consists of features *is worn on top of a T-shirt*; *has long sleeves*; *has sleeves*; *has two sleeves* that mainly describes about concept-cluster 5 but also some items in concept-clusters 3 and 4 for all the four subjects. Similarly, there are strong links between f8 with the concept cluster 6; f10 with the concept-cluster 4 and so on. Another tendency identified from Figure 5 is that the feature-clusters (e.g. f8,

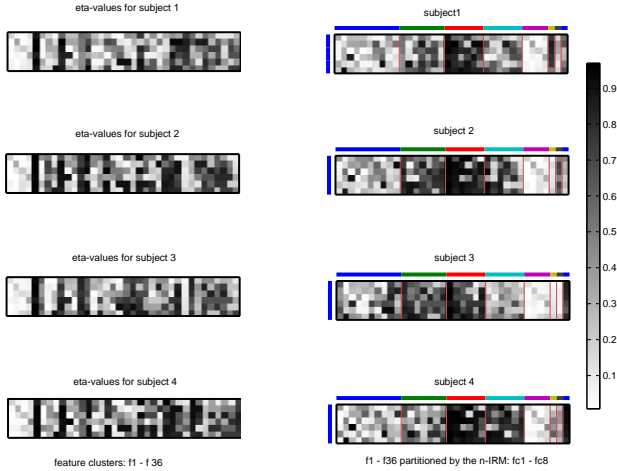


Figure 3: Individual semantic structures contrasted by η -values: left - feature clusters are extracted by the IRM; right - feature clusters are partitioned by the n-IRM

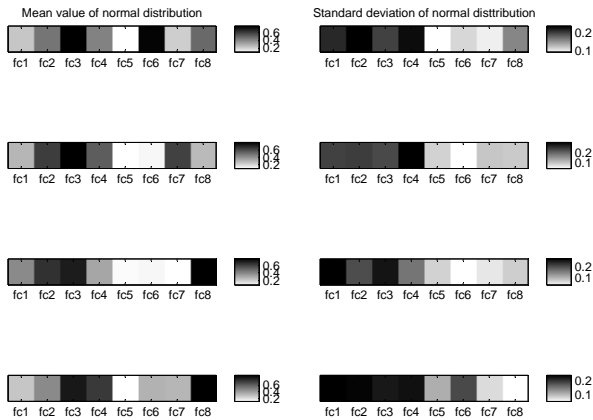


Figure 4: Mean values and standard deviations for the feature-cluster clusters (fc) extracted by the n-IRM across the four subjects

f18) that consist of features describing physical characteristics, e.g. *has sleeves, has zipper* etc. have similar patterns of associations possessed by the four subjects, while the feature-clusters (e.g. f15, f24) that consist of features describing impressions, e.g. *is feminine, can be a status symbol, fashionable* are rather subject-dependent.

In Figure 4, the mean values for *fc2* are generally higher than those of *fc1* for all the four subjects. This indicates that the feature-clusters are linked with several concept-clusters. At the same time, both the mean values and the variance indicate that the pattern of feature associations to the concept-clusters are different across the four subjects. For example, the feature-cluster f19 (is used to keep you warm; gives warmth; is warm; is warm in winter) in *fc2* clearly indicates that the patterns of association to the concept-clusters are different especially for the subject 3 in 6. Most of the features included in *fc2* describe functional-

ity such as *is worn to keep you warm* and impression such as *worn by elderly, is loose* etc.

The mean values for *fc3* are all very high in Figure 4, while there are some slight differences in variances across the four subjects. The high mean values indicate that each feature-cluster contained in *fc3* has tendency to form stronger links with majority of the concept clusters. In other words, the features listed in these feature-clusters, e.g. *bought in a store, exists in different kinds* etc., are rather common to most of concepts so that they have little influence on the formation of the concept-clusters.

As it can be observed from both Figures 4 and 6, *fc4*, *fc6*, *fc7* and *fc8* respectively consist of the feature-clusters that are rather subject-dependent. For example, subjects 2 and 4 showed different association patterns for features such as *is comfortable, is fun to wear, is sexy* from the patterns indicated by subjects 1 and 3. On the contrary, for *fc7* consisting of features such as *is hygienic, is beautiful, and smells like washing powder*, subjects 1 and 4 indicated similar patterns, while subject 2 and subject 3 showed completely different association patterns in 6.

Finally, Figure 7 represents an overview of *fc5* that consists of very sparse feature-clusters, i.e. f1, f2, f3 and f4. These feature-clusters consist of features that describe specific clothing items rather than the all items in a specific concept-cluster. For example, the feature *keep your feet warm* in f1 specifically associates with the concepts *boots, shoes* and *socks* that are members of the concept-cluster 1, while the feature *keep your head warm* in f1 specifically associate with the concepts such as *beanie, hat, scarf* and *cap* in the same concept-cluster 1.

The approach of applying the IRM to the multi-array matrices enables one to first identify a common semantic structural grid across multiple human subjects, from which different patterns of semantic structures possessed among the human subjects can be analyzed in details. By applying the n-IRM to the four η -value matrices obtained from the IRM, it is further possible to roughly estimate hierarchical structure of 29 clothing items via the item-specific features shown in *fc5*, the concept-cluster specific features shown in *fc1*, and the features that cover multiple concept-clusters such as *fc2*. This indicates an interesting potential for learning ontological structures of a collective domain knowledge possessed by multiple people. Future perspectives of this approach are further discussed in the next section.

5. Discussion and Future Perspectives

The work presented in the previous sections demonstrates that semantic representations of domain knowledge are slightly different across the four individual human subjects. For example, subject 2 and 4 indicated similar pattern of association to a specific feature-cluster, while, for another feature-cluster, the same subject 2 indicated different pattern of association from subject 4 but similar pattern possessed by subject 3. One of the main limitation in this work is that the publicly available datasets made by [De Deyne 08] do not disclose the profiles of each subject.

Hence it is impossible to investigate if the identified differences in the association patterns are caused by aspects of the subjects' profiles such as gender.

One of our future plans is to collect data from a larger number of human subjects with different profiles, such as cultural backgrounds, genders, generations, social environments, different expertise etc. To be more specific, a group at Copenhagen Business School is planning to study patterns of semantic structures of a specific knowledge domain, possessed by 150 subjects in 6 countries, i.e. Denmark, Japan, China, Russia, Spain and UK, and investigate how a conceptualization differs according to the different types of mother languages [Durst-Andersen 11], cultural backgrounds, social status, gender, English proficiency level, etc., when domain-specific terms are expressed in a common language, i.e. English. The work demonstrated here indicates that the application of the IRM seems to be an interesting starting point for analyzing such semantic structures of people having different profiles. The results presented above also imply some exciting future potentials for: i) the unsupervised construction of hierarchical ontologies of the domain knowledge based on the analysis of the feature-concept structure; ii) the clustering of 150 subjects based on different patterns of concept-feature relations across 150 matrices; iii) the identification of a collective ontology across subjects belonging to a specific profile group; and iv) the alignment of the identified collective ontologies that belong to two respective profile groups. These attempts could eventually contribute to the bottom-up learning of collective knowledge which optimally represents individuals' knowledge naturally possessed by a group of people.

6. Conclusions

The work introduced an approach of combining the IRM and the n-IRM for optimally identifying a common semantic structural grid from multiple concept-feature matrices, from which the individual semantic structures possessed by the respective subjects are further contrasted to each other. The results indicated that the IRM reasonably partitioned the 29 concepts into 6 clusters that respectively characterized with common themes. By further applying the n-IRM to the concept- and feature clusters obtained by the IRM, the feature-clusters are further structured in three levels: feature-clusters consisting of item specific features; feature-clusters consisting of a concept-cluster specific features; and feature-clusters consisting of features that cover several concept-clusters. In addition, the identified common grid effectively visualized the different feature association patterns possessed by the four subjects. The work indicated several potential models to be considered in the future research in order to achieve the bottom-up learning of collective knowledge which optimally represents individuals' knowledge naturally possessed by a group of people.

7. Acknowledgements

This work has been made as a preliminary study of a framework project, the Global English Communication project, headed by Per Durst-Andersen at Copenhagen Business School, which is funded by the Carlsberg Foundation.

References

- [Gruber 92] Gruber, T.: Toward principles for the design of ontologies used for knowledge sharing, In *International Journal Human-Computer Studies*, vol 43, pp.907-928 (1992)
- [Madsen 04] Madsen, B. N., Thomsen, H. E. and Vikner, C.: Principles of a system for terminological concept modelling, In *Proceedings of the 4th International Conference on Language Resources and Evaluation*, ELRA, pp.15-19 (2004)
- [Declerck 10] Declerck, T., Krieger, H. U., Thomas, S. M., Buitelaar, P., O'Riain, S., Wunner, T., Maguet, G., McCrae, J., Spohr, D., Montiel-Ponsoda, E.: Ontology-based multilingual access to financial reports for sharing business knowledge across Europe, In *Internal Financial Control Assessment Applying Multilingual Ontology Framework*, J. Rooz, and J. Iwanyos (Eds.) Budapest: HVG Press (2010)
- [Pierce 58] Peirce, C. S.: *Collected papers of Charles Sanders Peirce, I-VIII*. Cambridge: Cambridge University Press (1932-1958)
- [Durst-Andersen 11] Durst-Andersen, P.: *Linguistic Supertypes: A Cognitive-semiotic Theory of Human Communication*. Berlin-New York: De Gruyter Mouton (2011)
- [Sperber 86] Sperber, D., Wilson, D.: *Relevance: Communication and Cognition*. Oxford: Blackwell (1986)
- [Rosch 75] Rosch, E., and Mervis, C.B.: Family resemblances: studies in the internal structure of categories, In *Cognitive Psychology*, vol.7, pp.573-605 (1975)
- [Storms 97] Storms, G., and De Boeck, P.: Formal models for intracategorical structure that can be used for data-analysis, In K. Lamberts and D. Shanks (Eds.), *Knowledge, concepts, and categories*, London: UCL Press, pp.439-459 (1997)
- [Kemp 06] Kemp, C., Tenenbaum, J. B., Griffiths, T. L., Yamada, T. and Ueda, N.: Learning Systems of Concepts with an Infinite Relational Model, In *The Twenty-First National Conference on Artificial Intelligence* (2006)
- [Mørup 10] Mørup, M., Madsen, K. H., Dogonowski, A. M., Siebner, H. and Hansen, L. K.: Infinite Relational Modeling of Functional Connectivity in Resting State fMRI, In *Proceedings of Neural Information Processing Systems* (2010)

- [Herlau 12] Herlau, T., Mørup, M., Schmidt, M. N. and Hansen, L. K.: Modelling Dense Relational Data, In IEEE International Workshop on Machine Learning for Signal Processing (MLSP), Santander, Spain (2012)
- [De Deyne 08] De Deyne, S., Verheyen, S., Ameel, E., Vanpaemel, W., Dry, M. J., Voorspoels, W., and Storms, G.: Exemplar by feature applicability matrices and other Dutch normative data for semantic concepts, In Behavior Research Methods, vol.40(4), pp.1030-1048 (2008)
- [Pitman 02] Pitman J.: Combinatorial Stochastic Processes, In Notes for Saint Flour Summer School (2002)

		f6	f8	f10	f11	f14	f15	f17	f18	f20	f24
		protects against the cold'									
		protects against the rain'									
		protects against the sun'									
		is built by your grandma'									
		is used to protect you'									
		is not worn in the summer'									
		hung on hallway'									
		can be used outside'									
		can be worn as ornament'									
		can be hit by yourself'									
		used as accessory'									
		is worn on your legs'									
		has a button'									
		has a zipper'									
		has long (trousers)legs'									
		has (trousers)legs'									
		has two (trousers)legs'									
		has bags'									
		is a pair of trousers'									
		is long'									
		you can put a belt in it'									
		is worn to work'									
		is worn at special occasions'									
		gives a businesslike impression'									
		'is chic'									
		for fancy occasions'									
		used to look nicer'									
		worn by business people'									
		worn at parties'									
		worn on girls' legs'									
		worn to go skiing'									
		is worn by groomsmen'									
		is cheap'									
		is expensive'									
		can't be worn in the church'									
		is worn by cowboys'									
		is small'									
		is narrow'									
		is light around the body'									
		is blown away by the wind'									
		'is light'									
		'is in style'									
		'is fashionable'									
		'is hard to iron'									
		can be a status symbol'									
		'is given a present'									
		worn with a suit'									
		is worn under a sweater'									
		is refreshing in the summer'									
		'is airy'									
		'is made in a light fabric'									
		worn underneath other clothes'									
		is worn on top of a t-shirt'									
		has long sleeves'									
		has two sleeves'									
		can have sequins'									
		can be carried backwards'									
		can be manufactured by yourself'									
		shrinks when washed'									
		'is feminine'									
		worn by transvestites'									
		'men think it is sexy'									
	subject 1	hat'									
		boots'									
		beanie'									
		cap'									
		belt'									
		shoes'									
		scarf'									
		socks'									
		mittens'									
		bra'									
		dress'									
		skirt'									
		panties'									
		top'									
		bathing suit'									
		tracksuit'									
		pyjamas'									
		dungarees'									
		shorts'									
		t-shirt'									
		blouse'									
		tie'									
		shirt'									
		suit'									
		coat'									
		sweater'									
		pullover'									
		pants'									
		jeans'									
	subject 2	hat'									
		boots'									
		beanie'									
		cap'									
		belt'									
		shoes'									
		scarf'									
		socks'									
		mittens'									
		bra'									
		dress'									
		skirt'									
		panties'									
		top'									
		bathing suit'									
		tracksuit'									
		pyjamas'									
		dungarees'									
		shorts'									
		t-shirt'									
		blouse'									
		tie'									
		shirt'									
		suit'									
		coat'									
		sweater'									
		pullover'									
		pants'									
		jeans'									
	subject 3	hat'									
		boots'									
		beanie'									
		cap'									
		belt'									
		shoes'									
		scarf'									
		socks'									
		mittens'									
		bra'									
		dress'									
		skirt'									
		panties'									
		top'									
		bathing suit'									
		tracksuit'									
		pyjamas'									
		dungarees'									
		shorts'									
		t-shirt'									
		blouse'									
		tie'									
		shirt'									
		suit'									
		coat'									
		sweater'									
		pullover'									
		pants'									
		jeans'									
	subject 4	hat'									
		boots'									
		beanie'									
		cap'									
		belt'									
		shoes'									
		scarf'									
		socks'									
		mittens'									
		bra'									
		dress'									
		skirt'									
		panties'									
		top'									
		bathing suit'									
		tracksuit'									
		pyjamas'									
		dungarees'									
		shorts'									
		t-shirt'									
		blouse'									
		tie'									
		shirt'									
		suit'									

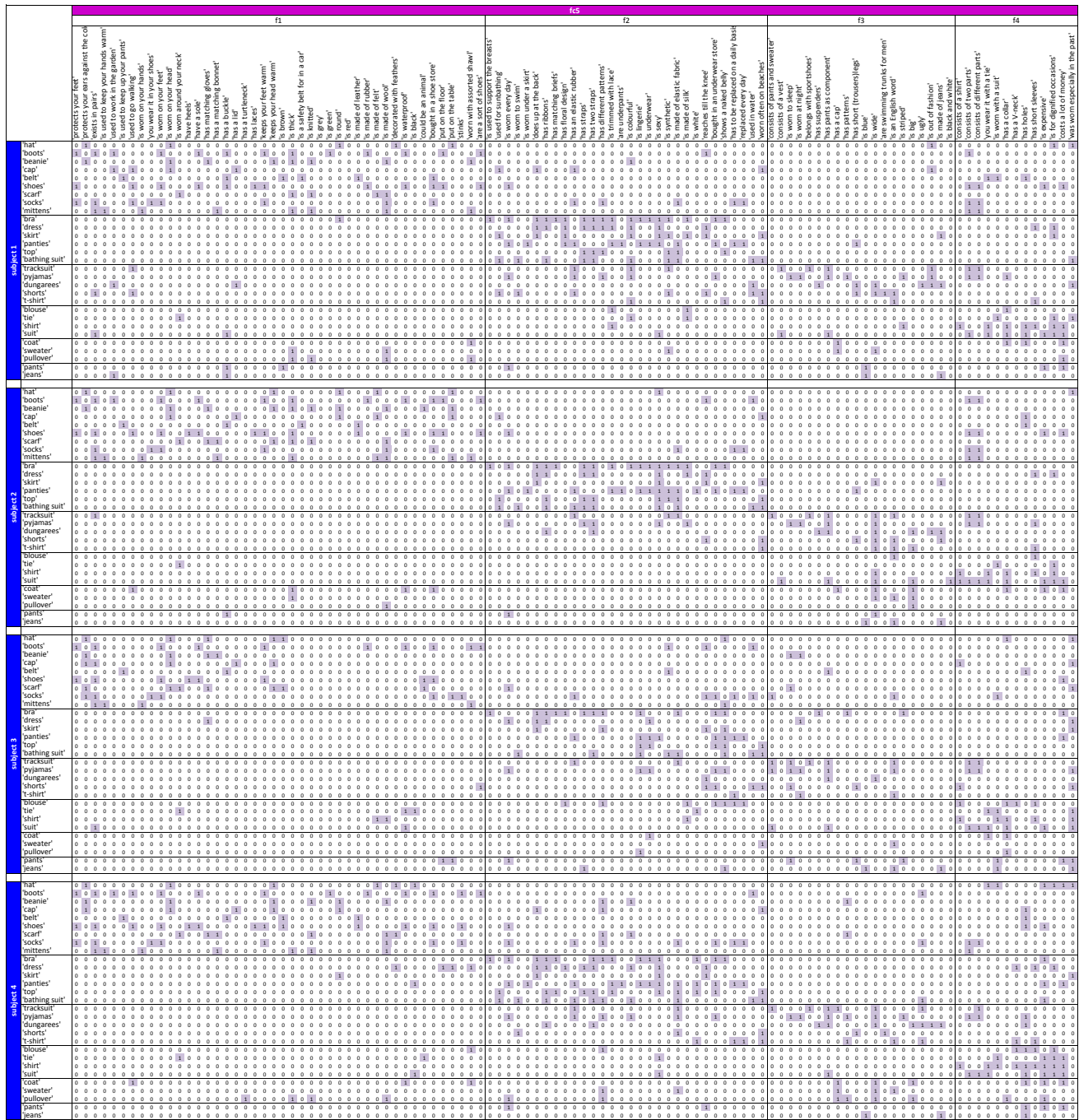


Figure 7: Obtained concept-clusters and feature-clusters for fc5