



A method, computer program and system for inferring and structuring relations between cultural specific concepts in two cultures

Glueckstad, Fumiko Kano; Moerup, Morten; Herlau, Tue; Schmidt, Mikkei, N

Publication date:
2013

Document Version
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Citation (APA):
Glueckstad, F. K., Moerup, M., Herlau, T., & Schmidt, Mikkei, N. (2013). A method, computer program and system for inferring and structuring relations between cultural specific concepts in two cultures. (Patent No. WO2013186216).

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



- (51) **International Patent Classification:**
G06F 17/28 (2006.01) *G06F 9/44* (2006.01)
G06F 17/30 (2006.01)
- (21) **International Application Number:**
PCT/EP2013/062040
- (22) **International Filing Date:**
11 June 2013 (11.06.2013)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
12171535.3 11 June 2012 (11.06.2012) EP
61/658,003 11 June 2012 (11.06.2012) US
- (71) **Applicants:** IBC - DEPARTMENT OF INTERNATIONAL BUSINESS COMMUNICATION [DK/DK]; Dalgas Have 15, DK-2000 Frederiksberg (DK). TECHNICAL UNIVERSITY OF DENMARK [DK/DK]; Anker Engelundsvej 1, Bygning 101A, DK-2800 Kgs. Lyngby (DK).
- (72) **Inventors:** GLÜCKSTAD, Fumiko Kano; N. Dalhoff's Vej 5, DK-2000 Frederiksberg (DK). MØRUP, Morten; Slotsvaenget 26, DK-2800 Kgs. Lyngby (DK). HERLAU, Tue; Århusgade 52, 1tv, DK-2100 Copenhagen Ø (DK). SCHMIDT, Mikkel, N.; Ved Kløvernmarken 12, st.tv., DK-2300 Copenhagen S (DK).
- (74) **Agent:** HØIBERG A/S; St. Kongensgade 59 A, DK-1264 Copenhagen K (DK).
- (81) **Designated States** (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (84) **Designated States** (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM,

[Continued on next page]

(54) **Title:** A METHOD, COMPUTER PROGRAM AND SYSTEM FOR INFERRING AND STRUCTURING RELATIONS BETWEEN CULTURAL SPECIFIC CONCEPTS IN TWO CULTURES

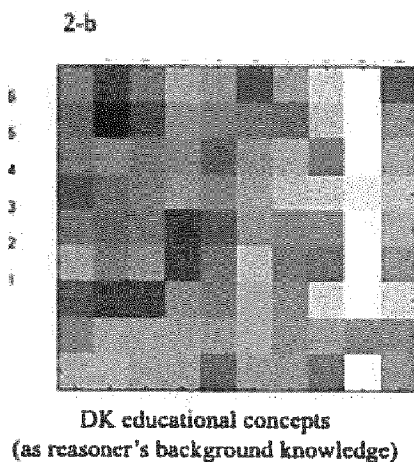


Fig 6

(57) **Abstract:** The present invention relates to a method, computer program and system for inferring relations between cultural specific concepts (CSC) in two cultures at least comprising the steps of - extracting and listing said cultural specific concepts (CSCs) and features of said CSCs from at least a first corpora belonging to a first culture and a second corpora belonging to a second culture, - applying a algorithm to infer relations between said CSCs in the first and the second corpora.

WO 2013/186216 A1



TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, **Published:**
KM, ML, MR, NE, SN, TD, TG).

— *with international search report (Art. 21(3))*

A method, computer program and system for inferring and structuring relations between cultural specific concepts in two cultures

5 **Field of invention**

The present invention relates to a method, system and computer program for inferring and structuring relations between cultural specific concepts (CSC) in at least two cultures.

10

Background of invention

Accelerated by the recent internet revolution with its fast-paced globalization, cross-cultural communication, e.g. between an Asian and a European, becomes inherently challenging due to the lack of sufficient linguistic resources directly bridging remote languages. Although we mostly use a common communication code such as English, misunderstandings are almost unavoidable in contemporary cross-cultural or cross-organizational communications. This challenge is not only caused by the lack of linguistic resources, but also by differences in human perception of similar concepts existing in diverse socio- cultural communities.

15

20

In many situations it becomes virtually an impossible task to precisely translate or convey the meaning of a Culturally-Specific Concept (CSC) if no exact equivalent concept exists in the Target Language (TL) culture.

25

Such challenge exists in many application domains. When it comes to the marketing activities, a product producer is supposed to convey the meaning of a new product concept to target consumers. How the new product is perceived, understood and positioned is determined by consumers' background knowledge. Here the challenge of cross-cultural communication between a producer and consumers exist. In the area of legal reasoning, for example the concept of "citizenship" may consist of a set of features such as "birth in Italy", "having Italian parents", "permission to stay in Italy", "right to vote" etc. in Italian legal system. Here, the legal concept "citizenship" plays as a mediator role and meanings of the concept is fully determined by the Italian legal norms. Hence if "citizenship" refers to the concepts determined by Japanese legal norms, not Italian norms, then the meanings of "citizenship" is differently interpreted between Japanese- and Italian reasoners.

30

35

The well-known Google translation is an interesting example of difficulties in existing approaches. When translating the following Japanese text into Danish, the Google solution obviously employs the pivot translation approach. Hence culturally-specific terms such as *Senmon-gakko* (official English translation: "Specialized training college, specialist degree course") which, according to actual bi-cultural knowledge, possibly corresponds to the Danish term "Erhvervsakademiuddannelse" (synonym: "kort videregående uddannelse") is translated into "college" or "school" by the Google translation. In the same way, *Koto Sensyu Gakko* (official English translation: "Specialized training school, upper secondary course") which corresponds to the Danish term "Erhvervsuddannelse" is translated into "Videregående erhvervsskole" via the pivot translation "higher vocational school". A critical problem lying here is that a solely statistical solution considers any kind of in this case English expression as one large bag of words (as English corpus) which contains not only expressions of concepts originating from English-speaking countries but also English translations of concepts originating from other non-English speaking countries. When considering the obvious goal of translation, that is, to convey the original meaning of a source concept to an audience in a target culture, to achieve the most successful cross-cultural knowledge transfer and, to share the common knowledge at the maximum degree between the two parties, a critical question is how well such a pivot translation can convey an original conceptual meaning of a source language (SL) word into a TL translation.

This problem could be well-explained by the following example: Imagine a situation where a Japanese diplomatic officer receives a business card from a Danish *konsulent* at *Økonomistyrelsen*. The Japanese officer needs to understand the definition of three words *konsulent*, *økonomi* and *styrelsen*, in order to figure out at which level this Danish person is positioned and in which kind of organization, by reflecting on his/her current conceptual understandings of the Japanese political system. First of all, a term like *Økonomistyrelsen*¹ consists of several words. In addition, each Danish word has more than one sense (meaning). In the case of *konsulent*, *økonomi* and *styrelsen*, each word respectively has one, five and four senses according to the Danish WordNet. When checking up with the available Danish-English dictionaries, there are several English translation options: ***konsulent*** can potentially be translated as *consultant*, *advisor*, *advisory officer*, *reader*, or *counselor*; ***økonomi*** into *finance*, *economy* or

economics; and **styrelsen** into *management, executive committee, council, board, steering committee, government agency, executive agency, state agency, and administration*. In addition to this, there are several Japanese translation options for each English translation candidate when looking them up in English-Japanese dictionaries. Since a term consists of several words and a word often carries several meanings, the dictionary-based transitive translation approach using English as a pivot language simply amplifies the probability of selecting an inappropriate sense in a TL. Thus the problem of word sense ambiguity becomes especially serious in the process of pivot translations.

In computer sciences, ontology has often been employed for defining domain knowledge for the purpose of achieving common understandings among members of a specific knowledge community. When it comes to the interaction across communities, i.e. cross-cultural communication, diverse methods for mapping ontologies have been introduced in recent years such as (Cheng et al., 2008), (Euzenat and Valtchev, 2004), (Ichise et al., 2004), (Ehrig, 2007), (Mitra et al., 2005). The traditional ontology mapping is based on the prerequisite that well-organized and hierarchically-structured domain specific ontologies exist. Accordingly, the focus of the ontology mapping is primary on the relevancy analysis i.e. similarity computation between concepts existing in two ontologies. The similarity computation usually employs algorithms that compute semantic distance between two concepts in question based on semantic information extracted from the existing ontologies. However, the relevancy analysis is not enough to identify a most relevant candidate corresponding pair and align knowledge across two ontologies. Hence the ontology alignment often requires database experts or knowledge engineers subjectively to manually align the hierarchically structured information. Another issue is that the strict logic rules applied for constructing an ontology before aligning the ontologies may cause risks for eliminating important information (feature attributes) that is necessary to compute relevancy between the two ontologies (Glückstad and Mørup 2012).

Another point is that a source concept often carries so-called equivalent expressions in several languages, i.e. an original expression in a source language and its official translation in English. A problem is that these lexically expressed multi-word expressions in different languages are sometimes semantically inconsistent. A typical example is illustrated by the case where the formal English name of the Danish

authority *Økonomistyrelsen* is *The Danish Agency for Governmental Management*.

When readers see these expressions without knowing the synonymous relationship between them, it is impossible for them to judge that these expressions refer to the

same concept. This example shows that it is a major challenge to find an appropriate

5 translation e.g. Japanese, that can optimally convey the original meaning of a specific Danish concept to a Japanese audience. One obvious point is that it is impossible to translate such a term without knowing its original conceptual meaning – that is, the definition of the concept under consideration. In the case of the Japanese translation of the Danish term *Økonomistyrelsen*, what readers need to know in a cross-cultural
10 communication context is the level at which the Danish organization is situated, information about whether the organization is part of the ministry, and what kind of authorization the organization has, etc. Hence, an ideal Japanese translation should reflect on these pieces of knowledge to the maximum extent, contrasted to the Japanese conceptual structure in question.

15

Despite this inherent difficulty, communicators or translators are still required to convey such CSCs into a Target Language (TL) in an optimal manner such that a TL reader can instantly infer the original meaning of a given Source Language (SL) concept.

20

This challenge of translating CSCs is not only caused by the absence of equivalent concepts in a TL culture, but also due to differences of the background knowledge possessed by the two parties involved in a cross-cultural communication scenario.

25

Thus, translation between not only two languages but also two different cultures contains special challenges which are not met by the traditional translation approach.

30

These traditional translation approaches are usually based on bilingual dictionaries (in a digital format) or statistically based approaches employing a large set of text corpora. A standardized terminology management where a universal (common) ontology plays as a central role and lexical expressions are localized into different languages as well as lexical semantic networks linking two concepts in a symmetric manner is also known.

35

Summary of invention

5 In a first aspect of the present invention is provided a method enabling consideration to cultural bias

In a second aspect of the present invention is provided a method enabling translation between different cultures which are not confined by language or national boundaries

10 In a third aspect of the present invention is provided a method facilitating cross-cultural conceptualization.

In a fourth aspect of the present invention is provided a method enabling structuring of the knowledge of one or more communicating parties.
15

These and further advantages are achieved by the present method for inferring relations between cultural specific concepts (CSC) in two cultures at least comprising the steps of
20 - extracting and listing said cultural specific concepts (CSCs) and features of said CSCs from at least a first corpus belonging to a first culture and a second corpus belonging to a second culture,
- applying an algorithm to infer relations between said CSCs in the first and the second corpora. By the present method it is possible, to obtain knowledge about relations
25 between cultural specific concepts in different cultures in order to evaluate links and similarities between one or more CSCs in one culture with one or more CSCs in another culture. The algorithm infers the relations based on the features defining the CSCs in each culture and thus the present method provides a way of linking culturally based concepts from different cultures in order to provide insight into e.g. how they
30 may be understood by a "receiver" rooted in a different culture.

Traditional tools for aligning knowledge possessed by two or more parties are generally based on ontologies that are the set of data structured based on ontological approach or lexical semantic networks approach.

35

Traditional tools for translations are generally based on digitally available linguistic resources such as lexicons, dictionaries, texts and audio visual data or statistical based translation using a large set of text corpora.

- 5 Traditional tools for aligning ontologies are generally using lexical semantic networks that consider a link between two concepts in different cultures as symmetrical relation.

10 In contrast to these traditional approaches the present method provides a way to acknowledge different cultural background and the meaning this will have for how a CSC from a first culture is best transferred (or aligned) with a CSC from a second culture. This is done by applying algorithms which are constructed to infer relations between CSC defined by features which algorithms are designed to encompass cognitive processes and/or back ground knowledge relevant for communicator and/or receiver.

15 The meaning of cultural background can be illustrated for example by looking at the concept "Energy". Energy is a word known to us all, but it is also a word with many definitions depending on the context it is used. For example Energy may have a different meaning to a person with a cultural background rooted in high energy physics compared to that of a person rooted in a culture holistic therapy even though they may be sharing a common language. This means that successful communication of what is meant by "Energy" may possibly be requiring use of other concepts or additional explanation.

25 The present method can provide insight into when care is to be taken during communication in order to avoid misunderstandings. Also the present method may suggest alternative concepts which may communicate the meaning better to a receiver.

30 As indicated in the example above, culture is not limited to national or regional culture but is seen broadly and also covers cultures of e.g. organizations whether private or public, local or international, of religious groupings, sub cultures relating to music or life style, regional, education, age etc.

The cultural specific concepts (CSCs) may relate to any culturally defined area, for example CSCs may be concepts known from the education system in two different countries, be concepts from different musical genres or to local dishes and receipts.

- 5 Each CSC is defined or described by a number of features. For example Danish “smørebrød” (smorgasbord) may be considered a CSC defined by the features: Lunch dish, cold food, various individual topping, rye bread, traditional food. The CSC “smørebrød” may be part of a group or cluster of CSCs relating to e.g. “traditional Danish food”.

10

In linguistics, a corpus (plural *corpora*) or text corpus is a large and structured, semi-structured and/or unstructured set of texts (now usually electronically stored and processed). They are used to do statistical analysis and hypothesis testing, checking occurrences or validating linguistic rules on a specific universe. In the present context a

15 corpus may be culture specific, for example relating to information and understanding of religious, professional, regional groupings.

From these corpora CSCs and their features may be extracted.

- 20 A corpus may here be represented in the form of a text, documents, compilation, diagram, hierarchical structure, audio, audio-visual data etc. from which the CSCs and features may be extracted.

- The first and second corpus may be a “merged corpus” i.e. may be part of or defined
- 25 by the same resource such as text, database, audio and/or visual recourse.

Depending on how the corpus is represented the extraction may be performed manually, which for example may be required if the CSCs are to be extracted from one or more texts with none or limited assessable structure of the data.

30

Automatic extraction may also be relevant, for example in cases where the CSCs and features are extracted from a structured source corpus such as tables, ontologies or node structures. Other possibilities where automatic extraction of CSCs and features and also identification of corpora may be relevant is in relation to various web

35 applications and resources. I.e. other possibilities are to use tools for automatically

extraction or semi-automatically selecting relevant domain knowledge from web-resources and compiling corpus consisting of CSCs and their features¹.

5 Thus, extraction of the CSC may be performed manually, semi-automatically or automatically from various types of sources.

The extracted information of CSCs and features may be listed in several ways but for example it can be represented in tables (or matrices) as in fig 2a where as an example Japanese CSCs (J CSC) and there defining features are listed. For example the CSC
10 J1 possesses the features 2 and 3, but not features 1 and 4.

A similar table could be made to represent the CSC and features extracted from the second corpus e.g. belonging to a Danish corpus fig 2b.

15 From here a matrix (fig. 2c) representing the use of the algorithm to infer relations between the JP and Danish CSC can be made.

Depending on the type of algorithm used different values (a-i in fig. 2c) representing the connection or relation between the different Danish and Japanese CSC are achieved.
20

If a cognitive environment (culture) is shared by two parties, the set of all facts is manifest to both communicator and receiver and therefore this may possibly generate a common ground based on the symmetric coordination i.e. there is a 1:1 understanding of CSC as will be described later. However, from a realistic viewpoint, people use
25 different languages and are mastering different concepts as they are rooted in different cultures, so that the way people construct mental representations (organize taxonomies of domain knowledge) and make inference (interpret and/or understand e.g. a CSCs) are inherently different. A shared cognitive environment can also be described as the two parties having identical background knowledge.

30 By the present method it is possible to align Culturally-Specific Concepts (CSCs) that are respectively rooted in two remote languages and/or cultures. The idea of focusing on these CSCs is brought about by the following reasons: When focusing on problems

of cross-cultural communication occurring at the lexical level: **1)** often, it is impossible to identify a semantically 100% equivalent culturally-specific concept for a remote second culture; and **2)** prior-knowledge of culturally-specific concepts that is deeply rooted in one's own cultural background such as traditions, social systems etc.

5 inherently influences as cultural bias on any cross cultural communication scenario i.e. the way we understand new or foreign concepts depends on what we know.

These difficulties are addressed by the present method making it possible to improve the transfer of original conceptual meanings of SL concepts to a TL audience when
10 there are no 100% equivalent concepts existing between the two remote cultures under consideration and when both a TL audience and an SL communicator process inference based on their own culturally-rooted prior-knowledge.

The present method may also be referred to as Cognitive Ontology Mapping (COM) or
15 alternatively Cross Categorization Approach (CCA). When the term COM is used CCA also may apply and vice versa.

Cognitive Ontology Mapping is to identify corresponding concepts existing in two remote cultures. The scope of the concept mapping is preferably limited to culturally-
20 specific domain knowledge at the level of individual lexical items – not the entire utterance. Hence general vocabularies such as verbs and adjectives etc. are generally not considered by the present method. However, the present invention may be applied in various ways on various parts of or types of utterances and vocabularies.

25 COM is to identify corresponding concepts and/or categories (taxonomically structured classes) existing in two cultures. The approach taken here is the terminological approach, which considers that a concept is having at least one or more linguistic expressions. For one concept, it is allowed to possess several expressions.

30 If CSCs from different languages are to be related by the present method a pivot (intermediate) language may be chosen and at least the cultural specific concepts and features in the at least first and second corpora may be translated to the pivot language. This is for example relevant when concepts from Japanese and Danish are treated by the present method. In this case the Japanese CSCs with their belonging features and
35 the Danish CSCs with belonging features are both translated into e.g. English which

then is the pivot language. However, other languages than English can be used e.g. even a meta-language used for semantic annotation can be used. The lists or matrixes of CSC and features may then be created in the pivot language instead of having one matrix in e.g. Japanese and one in Danish.

5

In the present application Japanese and Danish, or Danish and German are often used as examples of different cultures and different spoken and written languages. However the present method may be applied to any combination of cultures and languages, including mono-lingual different cultures.

10

The COM/CCA aligns and structures semi-structured or unstructured datasets existing in two preferably heterogeneous domain knowledge (e.g. the educational systems belonging to two different legal districts, Japan and Denmark).

The workflow of COM accommodates algorithms for inferring relations between CSCs either in a symmetric or a asymmetric manner.

15

The workflow of COM also accommodates algorithms for identifying latent ontological (hierarchical) structures from two independent semi- or unstructured data sources while inferring relations between the CSCs in the two independent data sources. Hence the method of this invention is distinguished from the traditional ontology alignment approaches which primarily address issues on integrating already existing ontologies.

20

Preferably the algorithm (i.e. for inferring relations between CSCs) is a Bayesian inference model whereby the present method is employing cognitive models as algorithms to computationally map culture specific concepts and infer relations between CSC based on a cognitive approach. I.e. the algorithm for inferring relations between CSCs is a model whereby features possessed only by a receiver is considered and the importance of features are weighted based on the number of concepts possessing a specific feature in question.

25

30

Such a model may be a Bayesian model of generalization (BMG) (Tenenbaum & Griffiths, 2001).

A Bayesian model of generalization (BMG) can be described as follows:

35

Tenenbaum and Griffiths (2001) argue that Tversky's Ratio Model (Tversky, 1977) which compute similarity between a concept x and y based on common features and distinctive features possessed by the respective concepts, is remarkably similar to the Bayesian Model of Generalization (BMG), that is rooted in Shepard's theory of the generalization problem. Hence, it is possible to unify these two classically opposing approaches to similarity and generalization. In (Tenenbaum and Griffiths, 2001), three crucial questions of learning, after Chomsky 1986, are addressed in order to explain the BMG: **1) what constitutes the learner's knowledge about the consequential region;** **2) how the learner uses that knowledge to decide how to generalize;** and **3) how the learner can acquire that knowledge from the example encountered.** For instance, one example x of some consequence R is given. It is assumed that x can be represented as a point in a continuous metric psychological space and R corresponds to some region (referred to as *consequential region*) of this space. A task of the learner is to infer the probability that a newly encountered object y will fall within R given the observation of the example x . This conditional probability can be expressed as: $P(y \in R|x)$. In order to compute the conditional probability, Tenenbaum and Griffiths (2001) first answer the question: the learner's knowledge about the consequential region is represented as a probability distribution $p(h|x)$ over an a priori-specified hypothesis space H of possible consequential regions $h \in H$. Prior to observing x , this distribution for the prior probability is $p(h)$, then becomes the posterior probability $p(h|x)$, after observing x . According to (Tenenbaum and Griffiths, 2001), the learner uses this knowledge for generalization by summing the probabilities $p(h|x)$ of all hypothesized consequential regions that contain y as follows

$$P(y \in R|x) = \sum_{h: y \in h} p(h|x) \quad (1)$$

25

Tenenbaum and Griffiths (2001) further describe how a rational learner arrives at $p(h|x)$ from $p(h)$ after the generalization, through the use of Bayes' rule as follows

$$P(h|x) = \frac{p(x|h)p(h)}{p(x)} \quad (2)$$

Tenenbaum & Griffiths (2001) argue that equation (1) can be similar to Tversky's Ratio Model. In order to demonstrate this, they have reformulated Tversky's Ratio Model to the following formula:

$$sim(y, x) = 1 / \left[1 + \frac{\alpha \cdot f(Y - X) + \beta \cdot f(X - Y)}{f(Y \cap X)} \right] \quad (3)$$

- 5 As defined in the previous section, X and Y are the feature sets representing x and y , respectively. f denotes a measure over the feature sets, here considered as an additive function. $(Y \cap X)$ represents the sets of features present in both Y and X , $(Y - X)$ represents the sets of features present in Y but not in X , and $(X - Y)$ represents the sets of features present in X but not in Y . α and β are free parameters representing an asymmetric relationship between Y and X . A key point in this invention is to clarify which variable is defined as concept in an SL- or a TL culture. According to Tversky (1977), *if $sim(y, x)$ is interpreted as the degree to which y is similar to x , then y is the subject to the comparison and x is the referent*. Considering that *similarity serves as an organizing principle by which individuals classify objects, form concepts, and make generalizations* (Tversky, 1977), the most similar concept to an SL concept which is correctly identified in a TL culture of an audience through a feature mapping could, from a communicator's point of view, be the set of assumptions which are adequately relevant to the audience. Hence, the referent x should be defined as an SL concept and the subject y , that is to be compared with x , should be defined as a TL concept according to Tversky's Ratio Model. This definition should be applied to all algorithms derived from Tversky's Ratio Model.

Keeping this definition in mind, Tenenbaum and Griffiths (2001) have reformulated equation (1) in a mathematically equivalent form as follows:

$$P(y \in R|x) = 1 / \left[1 + \frac{\sum_{h: x \in h, y \notin h} p(h, x)}{\sum_{h: x, y \in h} p(h, x)} \right] \quad (4)$$

- 25 A key point emphasized by Tenenbaum & Griffiths (2001) here is that *the bottom sum ranges over all hypotheses that include both x and y , while the top sum ranges over only those hypotheses that include x but not y . If we identify each feature k in Tversky's framework with a hypothesized subset h , where an object belongs to h if and only if it possesses feature k , and if we make the standard assumption that the measure f is*

additive, then the Bayesian model as expressed in equation (4) corresponds formally to Tversky's Ratio Model (3) with asymmetric parameters $\alpha=0$, $\beta=1$.

It means that if the free parameters in equation (3) is set as $\alpha=0$, $\beta=1$, this algorithm is formally corresponding to equation (4) of the BMG which compute *the conditional probability that y falls under R (Consequential Region) given the observation of the example x* (Tenenbaum & Griffiths, 2001). The consequential region R , in this application, indicates the categorical region where a subject y belongs. In equation (4), a hypothesized subset h is defined as the region where a concept belongs to h , if and only if, it possesses feature k (Tenenbaum & Griffiths, 2001). Hence, in the COM framework, y has been implicitly considered as a newly encountered object existing in the TL ontology that should be aligned to the referent ontology that is to be considered as the background knowledge of a communicator in the SL culture. It means that by exchanging assignment of variables x and y , the algorithm defined in equation (4) also computes the probabilities that the audience in a TL culture based on the background knowledge of the audience generalizes an SL concept from a stimulus presented by a communicator.

Another key point of the BMG is that $P(h, x) = P(x|h)P(h)$ in equation (4) represents the weight assigned to the consequential subset h in terms of the example x . This can be achieved by specifically assigning the weight $P(h, x)$ based on the "strong sampling scheme" defined in (Tenenbaum & Griffiths, 2001) as follows:

$$P(x|h) = \begin{cases} 1/|h| & \text{if } x \in h \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

Here, $|h|$ indicates the size of the region h (Tenenbaum & Griffiths, 2001). According to Tenenbaum & Griffiths (2001), *what likelihood function $P(x|h)$ is determined by how we think the process that generated the example x relates to the true consequential region for R* . For example, Shepard's Universal Law of Generalization (1986), Heit's (1998) Bayesian analysis of inductive reasoning, and the standard machine learning literature (Haussler et al., 1994; Mitchell, 1997) argue that *the example x and consequential region R are sampled independently, and x just happens to land inside C* (Tenenbaum & Griffiths, 2001). Thus, the likelihood is defined in a binary fashion in the following way:

$$P(x|h) = \begin{cases} 1 & \text{if } x \in h \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

Opposed to this, Tenenbaum (1999) argues that *under many conditions, it is more natural to treat x as a random positive example of R , which involves the stronger assumption that x was explicitly sampled from R* . This argument leads to the “strong sampling scheme” defined by equation (5). In the COM framework, the number of objects possessing the k^{th} feature in the referent ontology, the background knowledge of either a communicator in the SL culture or an audience in the TL culture, is considered as the size of the region h . For example, if the feature “compulsory education” is possessed by three objects, the weight is assigned as 1/3. This strong sampling can intuitively be illustrated in a situation where the feature “objects that have four legs” is given to us as an example. We immediately imagine that this object must be something related to an animal or possibly a piece of furniture. We unconsciously limit the hypothetical region within a narrower region in order to achieve a more effective generalization. Finally, Tenenbaum & Griffiths (2001) explain that the prior $P(h)$ is not constrained in their analysis so that *it can accommodate arbitrary flexibility across contexts*. Hence, in this application, we set $P(h) = 1$ but this is not to be construed as limiting to the invention. However, strictly speaking, $P(h)$ could be computed based on probabilities that a hypothesized subset h is identified among all possible hypothesized subsets h that are assumed from a dataset in question.

Preferably the algorithm puts a bias on background knowledge and/or the significance of a feature is reversely proportional with the occurrence of said feature i.e. the algorithm can be selected to encompass that a receiver's interpretation of a CSC is biased by what is already known to this person based on his or hers cultural background. Thus a receiver being presented with a new or partly new concept will interpret this new concept based on prior knowledge and this may preferably be recognised by the algorithm.

Bias here is a learner's (either a communicator or a receiver) background knowledge expressed as feature sets of each categorical concepts that are taxonomically represented as a dataset. Considering this datasets consisting of features and concepts possessing these features as background knowledge, the hypothetical space h is considered as a space where a specific feature k is belonging.

The algorithm can be designed to minimize the influence of features which occurs frequently as discussed above with the $1/h$ relation. This can be an advantage as features, which applies to all or many CSC in a group provides little or no information for selecting the best possible corresponding CSC.

5

As the algorithm can be biased to enhance the weight of features known by the person being present to a concept, the algorithm can be regarded as asymmetric which means that the outcome depends on from which point of view a CSC is being interpreted. This asymmetry means that different information is obtained if the Bayesian algorithm is applied from the first and/or second culture's point of view (fig. 20a and 20b).

10

Tversky's Rational model may also be used according to the present invention.

Tversky (1977) states that *similarity serves as an organizing principle by which individuals classify objects, form concepts, and make generalizations*. Tversky's view of similarity is distinguished from the traditional theoretical analysis c.f. (Shepard, 1987) on two points: **1)** while the theoretical analysis of similarity relations has been dominated by the continuous metric space models, Tversky argues that *the assessment of similarity between objects may be better described as a comparison of features rather than as the computation of metric distance between points*; and **2)** although *similarity has been viewed by both philosophers and psychologists as a prime example of a symmetric relation*, the asymmetric similarity relation has been demonstrated in (Tversky, 1977) based on several empirical evidences.

20

Based on these two points, Tversky proposed a classic set-theoretic model of similarity, later coined Tversky's Ratio Model, as described by the following equation:

25

$$sim(y, x) = \frac{f(Y \cap X)}{f(Y \cap X) + \alpha * f(Y - X) + \beta * f(X - Y)} \quad (7)$$

30

Here, X and Y are the feature sets of object x and object y , respectively. f denotes a measure over the feature sets. In this application, the function f is defined as additive in the series of empirical studies described in the next chapter. $(Y \cap X)$ represents the sets of features present in both X and Y , $(Y - X)$ represents the sets of features present in Y but not in X , and $(X - Y)$ represents the sets of features present in X but not in Y . Since the similarity score in equation (7) is normalized, the obtained score lies between

0 and 1. α and β are free parameters representing an asymmetric relationship between X and Y . Assignment of these parameters severely influences the similarity measurements. When defining $\alpha = \beta = 1$, $\text{sim}(y, x) = f(Y \cap X) / f(Y \cup X)$ corresponds to the well-known algorithm, Jaccard's coefficient measure (Jaccard, 1901). When

5 defining $\alpha = 1$ and $\beta = 0$, $\text{sim}(y, x) = f(Y \cap X) / f(Y)$ corresponds to what is found in e.g. (Bush & Mosteller, 1951). The uniqueness of Tversky's view, is this asymmetric similarity that has been originally demonstrated in (Tversky, 1977) based on several empirical evidences. His argument is that similarity does not necessarily form a symmetric relation. When $\alpha = \beta$, this similarity formula assesses the degree to which

10 object x and y are similar to each other. It means that $\text{sim}(y, x) = \text{sim}(x, y)$. On the other hand, if the parameters α and β differ, this symmetric relation does not hold. Tversky (1977) explains that *if $\text{sim}(y, x)$ is interpreted as the degree to which y is similar to x , then y is the subject of the comparison and x is the referent. Hence the features of the subject are weighted more heavily than the features of the referent (i.e., $\alpha > \beta$).*

15 *Consequently, similarity is reduced more by the distinctive feature of the subject than by the distinctive features of the referent.*

Tversky (1977) argues that his model *offers explications of similarity, prototypicality, and family resemblance* discussed in the previous chapter (Murphy, 2004; see also Rosch & Mervis, 1975). Based on his idea of typicality and asymmetric relations

20 between a subject and a referent, as well as on the Relevance Theory of Communication (Sperber & Wilson 1986) that inherits the asymmetric co-ordination between communicator and audience on the choice of code and context, translation should provide the set of assumptions that are adequately relevant to the audience. And the stimulus (that is translation) produced by the translator should be such that it

25 avoids gratuitous inferential processing effort on the audience's part. Considering that *similarity serves as an organizing principle by which individuals classify objects, form concepts, and make generalizations* (Tversky, 1977), the most similar concept to an SL concept, that is identified in the audience's taxonomic organization of categories through feature matching, could be the set of assumptions which are adequately

30 relevant to the audience. Based on this hypothesis, Tversky's Ratio Model assigning different combinations of α and β parameters is applied to datasets obtained based on the methods described herein.

In other embodiments the algorithm is based on parallel distributed processing, or the so-called connectionist approach, is an alternative to the Bayesian inference approach above. It has originally been developed as a neurally inspired model of the cognitive processes in human decision making. A connectionist network, like a biological brain, contains many highly interconnected, active processing units (like biological neurons) that communicate with each other by sending activation- or inhibition-signals through their connections. As in the brain, learning is implemented by modifying these connections, and knowledge is inherently represented in a distributed fashion over these connections. Specifically, computational implementations of a parallel distributed processing approach can be implemented in a plurality of ways but most typical implementation is by a so-called artificial neural network that is exposed to a plurality of training data in either a supervised way (where target values are known and expressing an error function to be minimized during training rounds) or an unsupervised way where gradually emerging data-clusters are formed based on these data examples.

15

For example a Danish translator who is looking for a relevant Japanese CSC will due to the bias on the "already known" evaluate the available Japanese CSCs based on his/hers Danish prior knowledge.

20

It is also possible to investigate what Danish CSC a Japanese audience would consider the best match to a given Japanese CSC, which possibly is not the same CSC as the initial Danish CSC.

25

Thus by the present invention it is possible to compute how CSCs are linked and understood depending on the culture of the person contemplating and aligning the CSCs.

30

Preferably the method comprises the step of identifying at least one candidate corresponding pair of a CSC from the first culture to the second culture i.e. the method can be used to find the best candidate(s) that a receiver will optimally infer about the meanings of the concept which the communicator intends to convey. As discussed above preferred algorithms are not necessarily symmetrical which means that the outcome may depend on the whether the link between the cultures are seen from the receiver or communicators point of view as the understanding of a CSC depends on what is already known by the person trying to comprehend a concept.

35

As the present method may be used to suggest what CSC in one culture may best provide the actual meaning of a concept (CSC) from another culture it is possible to suggest conceptual translations between languages or cultural backgrounds which otherwise would not have been realized.

Also the present method may include a step of identifying at least one probability that an information receiver belonging to the second culture (corpora) successfully infers the meaning of a CSC belonging to the first corpora translated to the second corpora, a step which may be applied with great advantage by cross cultural communicators. This step of analysing a translation of a CSC from the information receivers point of view overcomes the weakness which lies in the traditional translation approach which is based on the communicators ability to infer relations between CSC based on own knowledge only.

The method may also simply identify a symmetrical link between two concepts. Such algorithm can be based on among others Jaccard similarity coefficient which formula is the same as Tversky in the case $\alpha = \beta = 1$

In some situations it may be preferable if the method according to the present invention is further comprising the step of applying an unsupervised algorithm for learning taxonomies, and for structuring said hierarchical relations among said taxonomies and their features

If the CSCs and their features are extracted from a loosely or non-structured source it may not be clear which CSCs in a culture are related to each other and may be considered to form a group or class. The present method address this problem by applying a step wherein the elements in CSCxCSC (fig 2c) or CSCxfeature matrixes (fig 2b and 2c) are clustered in groups with common relevance by application of an unsupervised algorithm for learning taxonomies. From these clusters some order or taxonomy may be learned thereby providing a possibility to represent the CSCs from the two cultures in an ordered manner.

Such algorithms for clustering and imposing structure to data may be an infinite Relation model (IRM) which may be described as follows:

According to Kemp, et al. (2006), 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*. In (Kemp et al., 2006), the
 5 observed data are considered as m relations involving n types. For the experimental strategies **1)**, **2)** and **3)**, the simplest model: dealing with two types with a single two-place relation $R: T_1 \times T_2 \rightarrow \{0, 1\}$ is applied. More specifically, in strategies **1)** and **2)** T_1 corresponds to either Danish and/or Japanese CSCs and T_2 corresponds to features, while in strategy **3)**, T_1 and T_2 , respectively, corresponds to Danish CSCs and
 10 Japanese CSCs.

The principle of generating clusters in the IRM, according to (Kemp et al., 2006), is based on a distribution over partitions induced by a so-called Chinese Restaurant Process (CRP) (Pitman, 2002). The CRP starts a partition process with a single cluster
 15 containing a single object. The i^{th} object has possibilities to belong to either of the following:

- A new cluster with probability: $\gamma / (i-1 + \gamma)$
- An existing cluster with probability: $n_a / (i-1 + \gamma)$

20

Here, n_a is the number of objects already assigned to cluster a , and γ is a parameter (Kemp et al., 2006). The CRP continues until all the objects belong to clusters. Hence, the distribution over clusters for object i conditioned on the cluster assignments of objects 1, ..., $i-1$ is defined as (Kemp et al., 2006):

25

$$P(z_i = a | z_1, \dots, z_{i-1}) = \begin{cases} \gamma / (i-1 + \gamma) & a \text{ is a new cluster} \\ n_a / (i-1 + \gamma) & n_a > 0 \end{cases} \quad (8)$$

30

(Kemp et al., 2006) explains that *the distribution on z induced by the CRP is exchangeable: the order in which objects are assigned to clusters can be permuted without changing the probability of the resulting partition. $P(z)$ can therefore be computed by choosing an arbitrary ordering and multiplying conditional probabilities. Since new objects can always be assigned to new clusters, the IRM effectively has access to a countably infinite collection of clusters.*

From the clusters generated by the CRP, relations are generated based on the following generative model:

- As described above, for the cluster assignment of objects

$$z \mid \gamma \sim \text{CRP}(\gamma)$$

- For link probabilities between clusters

$$5 \quad \eta(a, b) \mid \beta \sim \text{Beta}(\beta, \beta)$$

- For links between objects

$$R(i, j) \mid z, \eta \sim \text{Bernoulli}(\eta(z_i, z_j))$$

10 The generative model is extended to handle two types/modes with a single two-place relation $R: T_1 \times T_2 \rightarrow \{0, 1\}$ by defining separate type/mode specific clusterings $z^{(1)} \mid \gamma^{(1)} \sim \text{CRP}(\gamma^{(1)})$, and $z^{(2)} \mid \gamma^{(2)} \sim \text{CRP}(\gamma^{(2)})$, such that $R(i, j) \mid z^{(1)}, z^{(2)}, \eta \sim \text{Bernoulli}(\eta(z_i^{(1)}, z_j^{(2)}))$.

15 In the above generative model, we set parameters $\beta=1$, and $\gamma^{(j)} = \log(J^{(j)})$ where $J^{(j)}$ is the number of CSCs in the j^{th} mode.

In here, relationships are assumed to be conditionally independent given cluster assignments (Kemp et al., 2006). The eventual purpose of the generative model is to identify a cluster z that maximizes $P(z \mid R)$. Based on the generative model defined above, relations from clusters are generated by:

20

$$P(R \mid z, \eta) = \prod_{ab} (\eta_{ab})^{m_{ab}^+} (1 - \eta_{ab})^{m_{ab}^-} \quad (9)$$

25 where m_{ab}^+ refers to the total number of links between categorical classes a and b ; and m_{ab}^- refers to the total number of non-links between categorical classes a and b . The conjugate prior η_{ab} is in the aforementioned generative model defined as:

$\eta(a, b) \mid \beta \sim \text{Beta}(\beta, \beta)$. Accordingly, the conjugate prior on η_{ab} is integrated out whereby the following is obtained:

$$P(R \mid z) = \prod_{a,b \in \mathbb{N}} \frac{\text{Beta}(m_{ab}^+ + \beta, m_{ab}^- + \beta)}{\text{Beta}(\beta, \beta)}, \quad (10)$$

Multiplying the distribution $P(z)$ induced by the CRP given in formula (8) to formula (10), we obtain the joint distribution $P(R,z)$ for which the posterior distribution $P(z|R) \propto P(R|z)P(z)$. According to (Kemp et al., 2006), the expected value of η_{ab} given z , is:

$$\frac{m_{ab}^+ + \beta}{m_{ab}^- + m_{ab}^+ + 2\beta} \quad (11)$$

5

The mathematical procedure for the inference is further described in Mørup M., Madsen K.H., Dogonowski A.M., Siebner H. and Hansen L.K (2010). Preferred solutions used are based on the sample with highest posterior distribution $P(z|R)$.

10

The Normal-Infinite Relational Model or n-IRM (Hansen T.J., Mørup M. and Hansen L. K., 2011), is a generalization of the aforementioned IRM to continuous data. The name normal IRM implies that the model employs normal observations and normal-inverse gamma priors instead of Bernoulli observation 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.

15

The elements of each $n_a n_b$ submatrix is parameterized by a mean intensity m_{ab} and its variance λ_{ab} . Accordingly, the generative model is defined for two types/modes as

20

$$\begin{aligned} \mathbf{z}^{(1)} &\sim \text{CRP}(Y^{(1)}) && \text{clustering of first mode,} \\ \mathbf{z}^{(2)} &\sim \text{CRP}(Y^{(2)}) && \text{clustering of second mode,} \\ \text{for all clusters } \lambda_{ab} &\sim \text{Gamma}(\alpha_0, \text{rate} = \beta_0) && \text{precision,} \\ \text{for all clusters } m_{ab} &\sim \text{Normal}(m_0, (\kappa_0 \lambda_{ab})^{-1}) && \text{mean,} \\ R_{ij} &\sim \text{Normal}(m_{z_i z_j}, \lambda_{z_i z_j}) && \text{links.} \end{aligned}$$

25

As with the IRM, it is possible to integrate out both the means and variances similar to what has been done in Equation 10 leading to an efficient sampler (see (Herlau et al., (2012) for further details). In equation (10), the prior values for the n-IRM are set to $\kappa_0=1$, $\alpha_0=15$ and $\beta_0^{-1}=10$ to reflect the scale of the similarity relations. For both the IRM and the n-IRM applications the number of iterations is e.g. chosen as 1000 where the first 500 are used for burnin.

30

Thus as briefly described in the previous sections, two types of relational models are preferably employed in our work. For the alignment of legal knowledge in two legal districts, we apply the normal/Infinite Relational Model (n-IRM) (Hansen T.J., Mørup M. and Hansen L. K, 2011) to conditional probabilities (similarities) computed by the BMG.

5 The n-IRM is supposed to co-cluster legal concepts existing in the two legal districts (cultures). After co-clustering the legal concepts in the two legal districts, the IRM is employed for uncovering relations between the obtained clusters of legal concepts and their features (inferential links).

10 The unsupervised learning algorithm can be applied prior to application of the mapping algorithm and/or after application of the mapping algorithm depending on the situation and the desired result.

If the IRM is applied prior to the BMG it is possible to analyse how features and each
15 categorical class relate to each other. Such information is useful for constructing feature-based ontology such as Terminological Ontologies (Madsen et al., 2004).

If the IRM is applied after application of BMG several advantages are achieved. For example direct application of the BMG enables a user to analyse further specific
20 similarity relations between category members of the respective categorical classes (groups) existing in the two cultures.

When applying the IRM after application of the BMG, scores representing similarity relations may be binarized in order to optimally obtain categorical classes (groups)
25 existing in the two cultures.

By applying the n-IRM instead of the IRM after application of the BMG, the categorization process can directly be achieved from the scores representing similarity relations because the n-IRM may handle any real numbers.

30

Also this may be the most effective approach for clustering CSCs into more specific and appropriate categorical classes as well as it may capture complex relationships existing between each categorical class in the two countries (cultures).

Even if the CSCs are not ordered with respect to each other the defining features can be listed for each CSC and the matrices known from fig. 2a – 2c can be created and the BMG can be applied to infer the relations between the various CSCs of the two compared cultures.

5

If IRM and/or n-IRM is applied to the unstructured CSCxCSC matrices clustering of linked or mutually relevant CSC can be imposed whereby the relations between CSCs may become obvious.

10

This combination of algorithms for clustering and imposing structure to data e.g. IRM and/or n-IRM and algorithms for inferring relations e.g. BMG according to the present method in fact may enable the construction of ordered representations of the CSCs belonging to the respective cultures but also to construction of common ordered representations of the CSCs in a manner which is meaningful from the point of view of both the first and second culture.

15

In several embodiments the method may comprise the step of describing differences in meanings between a candidate corresponding pair of a CSC from the first and the second cultures. I.e. the method may be used to enlighten the differences and possible difficulties in understanding between two cultures.

20

If the method comprises the step of constructing and/or visualizing at least one ontology, a simple overview of the obtained results can be formed. Especially the visualized ontology can be constructed by the unsupervised algorithm.

25

A preferred embodiment is a method for inferring relations between cultural specific concepts (CSC) in two cultures at least comprising the steps of :

- extracting and listing said cultural specific concepts (CSCs) and features of said CSCs from at least a first corpora belonging to a first culture and a second corpora belonging to a second culture,

30

- applying an algorithm to infer relations between said CSCs in the first and the second corpora where said inference algorithm is applied to unstructured or loosely structured data sets of cultural specific concepts and features to relate the CSCs of the first and second cultures,

- applying an unsupervised algorithm for learning taxonomies in said CSCs in the first and the second corpora, and for structuring relations among said taxonomies and their features, where said unsupervised learning algorithm is applied to structure the results of the algorithm to infer relations and wherein the result from the combined application of the said inference algorithm and the said unsupervised learning algorithm is used to construct at least one ontology.

An uniqueness of the present method is that the present COM/CCA method may address the problems of identifying latent ontological structures from two independent semi- or unstructured data sources while analysing/infering interactive relations between the respective domain knowledge in question. Hence the present approach is quite different from the traditional ontology mapping approaches which primarily address issues on integrating already existing ontologies.

This application, therefore, introduce an approach, i.e. COM/CCA , for aligning similar un- or semi-structured domain knowledge existing in two cultures for example existing in two heterogeneous legal systems such as the educational systems belonging to legal districts, Japan and Denmark.

As exemplified later it is possible to employ both symmetric similarity measures such as Jaccard similarity coefficient and generalization models such as the BMG for computing the degree of relations between all possible combinations of concepts existing in the two cultures (legal systems). To the obtained similarity scores, it is shown how to apply an extended version of the IRM, the so-called normal Infinite Relational Model (n-IRM) proposed by (Herlau et al., 2012) in order to cross-categorize the educational concepts existing in the two legal systems i.e. in two cultures.

The application of the IRM and/or n-IRM allows the user to identify number of categories, i.e. groups of e.g. educational concepts in the example, for the respective legal systems and the degree of the relations between categories in the two legal systems. Finally, the IRM may be applied to the original data consisting of legal concepts and their features (inferential links) for identifying underlying relationships, i.e. structures of inferential links, behind the specified concept system, i.e. the identified categories.

In this way, in some embodiments the method simultaneously categorizes CSCs e.g. legal concepts existing in two legal systems (cultures) and from there to structure two independent concept systems that are inter-operable in the most efficient manner. By employing the generalization model, e.g. the BMG, a reasoner B generalizes and
5 interprets a new legal concept A introduced from the legal district A by comparing the new concept A and his or her background knowledge of legal concepts B belonging to the legal system B.

The steps of the present method may be carried out in various order. For example:
10

a. Step-wise approach consists of three steps, a) concept mapping based on similarity measures, b) alignment of domain knowledge based on the n-IRM, and c) feature structure leaning by applying the IRM to the original data to the fixed clusters obtained from the n-IRM; or

15 b. One generic relational modeling approach where the system that can simultaneously align (cross-categorize), structure and draw hierarchical trees of domain knowledge possessed by two parties without going through the aforementioned three steps (BMG+n-IRM+IRM)

20 Thus if first inference algorithm is applied to unstructured data sets of cultural specific concepts and features to relate/align the CSCs of the first and second cultures, whereafter the unsupervised learning algorithm is applied to structure the results and the result from the combined application of the said inference algorithm and the said
25 unsupervised learning algorithm can be used to construct at least one ontology.

Such ontology may be feature-based ontology such as Terminological Ontologies (Madsen et al., 2004). In this way, the structured relations obtained from the present method may be visualized in a hierarchical tree that assists the identification of
30 concept-concept relations but also concept-category (superclass consists of several concepts), category-category, category-super category (more abstract category positioned in the higher level of a hierarchical tree) relations as well as the indication of specific semantic differences between the mapped concepts (corresponding pair of concepts/categories existing in two cultures).

35

Preferably the method comprises the step of identifying at least one candidate corresponding pair of a CSC from the first culture to the second culture i.e. the present method can be used to find the best candidate(s) that a receiver will optimally infer about the meanings of the concept which the communicator intends to convey. As discussed above preferred algorithms are not necessarily symmetrical which means that the outcome may depend on the whether the link between the cultures are seen from the receiver or communicators point of view as the understanding of a CSC depends on what is already known by the person trying to comprehend a concept.

As the present method may be used to suggest what CSC in one culture may best provide the actual meaning of a concept (CSC) from another culture it is possible to suggest conceptual translations between languages or cultural backgrounds which otherwise would not have been realized.

The present method can applied to the learning of an ontology from a semi- or unstructured data source belonging to one culture. This can for example be achieved by applying the inference algorithm to the same data source to compute similarity scores among concepts within the same data source.

The actual meaning of a concept (CSC) may be conveyed by referring an abstract concept (e.g. abstract category positioned in the higher level of a hierarchical tree) and listing features of the CSC which are not described by the abstract concept. This can be achieved by the present method.

Also the present method may include a step of identifying at least one probability that an information receiver belonging to the second culture (corpora) successfully infers the meaning of a CSC belonging to the first corpora translated to the second corpora, a step which may be applied with great advantage by cross cultural communicators. This step of analysing a translation of a CSC from the information receivers point of view overcomes the weakness which lies in the traditional translation approach which is based on the communicators ability to infer relations between CSC based on own knowledge only.

Accordingly is employed a model, e.g. the so-called Bayesian Model of Generalization (BMG) proposed by Tenenbaum and Griffiths (2001) for the analogical inference. The BMG computes a relation between e.g. two legal concepts based on characteristic features in a special way:

- 5 i) it considers features that are already known by a learner, e.g. a Danish learner who know nothing about the Japanese educational system cannot know all the characteristic features of the Japanese educational concept so that only features known by the Danish learner should be considered; and
- ii) it distinguishes importance of features and assigns weights to each feature
- 10 according to the degree of importance.

After computing the degree of relations between preferably all possible combination of concepts existing in the cultures such as two legal systems, can be applied an extended version of the IRM, the so-called normal Infinite Relational Model (n-IRM) proposed by (Herlau T., Mørup, M., Schmidt, M.N. and Hansen, L. K. (2012)) for cross-

15 categorizing the educational concepts existing in the two legal systems. The application of the n-IRM is applied in order to identify the number of categories, i.e. groups of educational concepts in this example, for the respective legal systems and the degree of the relations between categories in the two legal systems. Finally, the IRM is applied to the original data consisting of legal concepts and their characteristic features for

20 identifying underlying relationships, i.e. feature structures, behind the specified concept system, i.e. the identified categories.

Preferably the cultural specific concepts (CSCs) and features of the first culture are extracted from a first ontology and the cultural specific concepts (CSCs) and features

25 of the second culture are extracted from a second ontology

Extracting the CSCs and features from one or more ontologies may prove beneficial as the structure of the ontologies may help the extraction of the CSC and the features as well as the extracted CSCs may have a clear order with respect to each other, e.g. the

30 CSCs may be ordered in groups of related CSCs.

Also by choosing a specific ontology to extract the CSCs and features from it is possible thereby to choose a specific way of ordering e.g. the CSCs.

Specifically where the ontology is a terminological ontology (TO) Madsen et al 2004 the extraction of the CSCs with the belonging features may be performed with less hassle.

5 The possibilities within the COM framework and method are many. For example CSCs and features may be extracted from for example TOs i.e. from very structured data, or CSCs and features may be extracted from more loosely ordered(i.e. loosely structured) or unordered (i.e. unstructured) data. Algorithms, e.g. BMG or other of the algorithms discussed herein, can be applied to the extracted CSCs and features in order to infer relations between the CSCs.

10

In some cases, for example where the CSCs and features are obtained from loosely ordered datasets the application of BMG can advantageously be combined with IRM or other algorithms such as IRM and/or n-IRM for grouping the CSCs in clusters with. In fact the combination of BMG and IRM/n-IRM may be used in a process for generation of ordered structures such as TOs.

15

The present method can also be used for the automatic learning of an ontology in one culture from a semi-structured or unstructured data source. This can be achieved by applying the inference algorithm to the same data source to compute relations among concepts within the same data source.

20

Applications of the present method are many including but not limited to cross-cultural marketing, cross-cultural legal reasoning, expert-layman communication, organizational communication where the communications take place not necessarily between different languages but also between at least two parties possessing different background knowledge in the same language.

25

o In the marketing activities, the dataset could consist of e.g. products and their features (product specifications and other related features) in two parties, a producer and consumers.

30

o Further extension of the cross-cultural marketing application could be the cultural adaptation of product taxonomies used in e-businesses (online shops) operated across different cultures. Such function of cultural adaptation could be integrated as a culturally-specific recommender system that may also consider customers' profiles in a specific culture in question.

35

o In the expert-layman communication, the dataset could consist of e.g. medical knowledge (medical concepts and their features) possessed by health-care professionals and by patients

5 o In the cross-cultural legal reasoning, the dataset could consist of e.g. tax related concepts and their definitions and consequences existing in two or more different legal systems

10 o In the organizational communication, the dataset could consist of e.g. product knowledge (products and their features) possessed by different job-functions in a production line such as purchasing dept., production dept., R&D dept., sales dept. and marketing dept.

o The final application domain is the machine translation and cross-lingual information access where the existing natural language processing technologies are combined with the proposed system

15

The present invention also relates to a computer program for executing a method for inferring relations between cultural specific concepts (CSC) in two cultures at least comprising the steps of

20 - extracting and/or listing said cultural specific concepts (CSCs) and features of said CSCs from at least a first corpora belonging to a first culture and a second corpora belonging to a second culture,
- applying a algorithm to infer relations between said CSCs in the first and the second corpora.

25 Preferably the method implemented by the computer program is a method according to the present invention.

30 For example the computer program is a program for executing a method for inferring relations between cultural specific concepts (CSC) in two cultures at least comprising the steps of

- extracting and/or listing said cultural specific concepts (CSCs) and features of said CSCs from at least a first corpora belonging to a first culture and a second corpora belonging to a second culture,
- applying an algorithm to infer relations between said CSCs in the first and the second
35 corpora where said inference algorithm is applied to unstructured or loosely structured

data sets of cultural specific concepts and features to align the CSCs of the first and second cultures,

- applying an unsupervised algorithm for learning taxonomies in said CSCs in the first and the second corpora, and for structuring relations among said taxonomies and their features, where said unsupervised learning algorithm is applied to structure the results and wherein the result from the combined application of the said inference algorithm and the said unsupervised learning algorithm is used to construct at least one ontology.

The computer program may consist of one or more program modules which alone or together execute one or more of the steps in the method according to the present invention.

The computer program may consist of locally stored and/or remotely stored program modules, as well as the computer program may consist at least partly of internet and/or cloud based program modules.

The computer program may be an integrated or supplemental part to larger software frameworks.

The present invention further relates to a system comprising one or more devices comprising a computer program and/or program modules for enabling inferring relations between CSCs from a first and a second culture.

Devices includes but are not limited to, PCs, tablets and other handheld portable devices, local or remote resources as well as both standard devices and/or specially developed or refitted devices.

The system may also comprise devices for obtaining information from analogue, digital, written and/or audio sources. The system may also comprise devices for interacting with one or more users receiving, sending and/or processing information.

Preferably the computer program of the system which infers the relations between CSCs from a first and a second culture is a computer program according to the present invention.

The present method, computer program and system may be applied widely to enhance translations, understanding and/or infer relations between CSC in real time interactions between people or as a resource for individuals to obtain new and otherwise not obtainable insight.

5

Thus by the present invention is provided a method, computer programme and system each enabling translations of concepts between not only two remote languages but also between two different cultural understandings. One important feature of the present method (COM/CCA approach), computer programme and system is to apply mathematical models derived from cognitive science to the issue having been considered as one of the most challenging topics within the research domain of Semantic-Web and multilinguality, that is, linking of culturally-specific concepts that are semantically inconsistent.

10

15

In the following the above will be further explained with reference to the figures. The explanations and figures are exemplary and are not to be construed as limiting to the invention. Throughout the examples, figures and description is referred to legal systems, in form of an educational systems in Japan and Denmark in relation to CSCs, features and cultures. However, this is exemplary and the examples, descriptions etc. may relate to various other cultures and CSCs as discussed above.

20

Fig. 1 illustrates how a concept may be understood differently in different cultures
Fig. 2a-2c illustrates CSC and feature matrices which can be created according to the present invention

25

Fig. 3a and 3b Flowchart of embodiments according to the present invention

Fig. 4 an exemplary dataset

Fig. 5 exemplary results of application of IRM

Fig. 6 plots in relation to fig. 5

Fig. 7 plots in relation to fig. 5

30

Fig. 8 co-clustering results

Fig. 9 intersections of cultural concept clusters

Fig. 10 Labels for IRM clustering

Fig. 11 co-clustering results

Fig. 12 intersections of cultural concept clusters

35

Fig. 13 Labels for IRM clustering

- Fig. 14 co-clustering results
- Fig. 15 intersections of cultural concept clusters
- Fig. 16 Labels for IRM clustering
- Fig. 17 Similarity relations
- 5 Fig. 18 Representation of Japanese legal concept system
- Fig. 19 Representation of Danish legal concept system
- Fig. 20a – 20d illustrates how translation candidates may be selected depending on prior knowledge
- Fig. 21a- 21c illustrates the asymmetric cross cultural communication
- 10 Fig 22 shows dimension specification of a concept (TO)
- Fig 23a, 23b and 23c shows flow diagrams for different applications of IRM and BMG
- Fig. 24a, 24b and 24c shows results for application of BMG+IRM
- Fig 25a and 25b shows diagrams representing traditional view on Relevance Theory of Communication and a revised view on Relevance Theory of Communication
- 15 Fig 26 shows some of the possibilities of the COM framework

Detailed description and examples

- 20 As discussed initially cross-cultural communication presents difficulties to the involved parts. As an illustration hereof fig. 1 tries to illustrate how the concept “tree” means creates a different picture to persons grounded in different cultures. It is these difficulties which are addressed by the present invention.
- 25 TO
- Terminological Ontology* (TO) is a domain-specific ontology used for knowledge sharing (Madsen et al. 2004), which normally is applied to terminology work, cf. for example (ISO, 2000). The unique points of the TO method that differentiate it from other types of ontologies are primarily its feature specifications and subdivision criteria
- 30 (Madsen et al., 2004). A feature specification consists of a feature dimension and its value. Thus, the representation of a whole concept is a feature structure, i.e. a set of feature specifications corresponding to the unique set of characteristics that constitutes a particular concept (Madsen et al., 2004). What it basically means is that a hierarchical structure of concepts - i.e. a TO - is shaped based on the feature
- 35 specifications and subdivision criteria. Terminological ontologists argue that concepts

are defined in a language-dependent context, and therefore a TO is language dependent. A TO is developed within a knowledge sharing community, then dynamically updated and validated. If it is necessary to share knowledge with other communities, TOs developed in different communities should be compared, aligned and merged, as needed. While the aforementioned two mainstream projects; MONNET (where a standardized ontology is translated into different language) and KYOTO (based on that culturally specific semantic network is linked via a universal ontology and that a link between two concepts in two cultures is symmetric), both deal with complex ontologies involving huge data-sets, TOs usually handle smaller amounts of concepts. Based on the dimensions defined by (Cimiano et al., 2010), TOs could be categorized as follows:

1. The TO method handles a *culturally influenced domain*;
2. Since a TO is independently developed within a knowledge sharing community, it is considered as an *independent ontology*; and
3. Since concepts are defined in a language-dependent context, a conceptual structure is completely adapted to a target community. It means that the TO method is considered as a *functional localization*.

This categorization of the TO is indeed identical to the KYOTO framework. One notable point is that TOs that are developed in different language communities should directly be compared, aligned and merged as needed, while the KYOTO framework employs a shared ontology to anchor culturally-specific ontologies and maintains a considerable degree of *interoperability through a mediator-ontology*. Therefore, it is arguable that, the KYOTO framework should rather be categorized as an *interoperable ontology* compared with the TO that maintains 100% independency. Another notable point that should be emphasized here is that a TO could potentially be a suitable tool for simulating cognitive processes explaining a real-life cross-cultural communication scenario. Since a TO maintains a hierarchical structure constructed based on conceptual features possessed by each concept, this enables one to compute feature-based semantic relatedness between two concepts, while the KYOTO framework allows one to compute semantic relatedness based on the distance-based measure. This difference will result in an argument that the feature-based measure enables one to compute either a symmetric- or an asymmetric relationship between two concepts,

while the distance-based measure is only limited to compute a symmetric relationship between two concepts.

Throughout the present application Danish, Japanese and German is used as examples of TL and SL and first and second cultures, and English is used as an example of pivot language. However, the present method can be applied to other languages and cultures as well as other languages than English can be used as pivot language.

Data creation from TO

From the TOs constructed in the previous section, structured feature sets that represent definitions of each concept are enlisted in a country-specific matrix. Feature values extracted from respective country-specific English corpora might express the same feature in different ways. These inconsistent expressions of a feature can be manually aligned. Examples of extracted CSCs (term) and features are given in Tables 1 and 2.

Table 1: Terms and features extracted from a TO (German)

ID	Term	Feature-values
G2	preschool education	{ISCED97, children & young, ISCED0}
G5	kindergärten	{ISCED97, children & young, ISCED0, child welfare, 3-6y.o.}
G7	schulkindergärten & vorklassen	{ISCED97, children & young, ISCED0, preparation}
G10	primary education	{ISCED97, children & young, ISCED1}
G11	primary school	{ISCED97, children & young, ISCED1, <6-10y.o.<}
G13	secondary education	{ISCED97, children & young, ISCED2+3}
G14	lower secondary level	{ISCED97, children & young, ISCED2+3, <10-16y.o.<}

G15	school offering one single course	{ISCED97, children & young, ISCED2+3, <10-16y.o.< , single}
G16	hauptschule	{ISCED97, children & young, ISCED2+3, <10-16y.o.< , single , general basic, 5-9 th grade}
G18	gymnasium	{ISCED97, children & young, ISCED2+3, <10-16y.o.< , single, intensified, 5-12/13 th grade}
G19	schools offering several courses	{ISCED97, children & young, ISCED2+3, <10-16y.o.< , several}

Table 2: Terms and features extracted from a TO (Denmark)

ID	Term	Feature-values
D2	pre primary	{ISCED97, children & young, ISCED0}
D4	kindergarten	{ISCED97, children & young, ISCED0, 3-6y.o.}
D6	single structure	{ISCED97, children & young, ISCED1+2}
D7	alternative structure	{ISCED97, children & young, ISCED1+2, alternative}
D9	efterskole or youth school	{ISCED97, children & young, ISCED1+2, alternative, compulsory, <14-18y.o.<}
D10	efterskole	{ISCED97, children & young, ISCED1+2, alternative, compulsory, <14-18y.o.<, boarding school, approved by state}
D11	youth school	{ISCED97, children & young, ISCED1+2, alternative, compulsory, <14-18y.o.<, day-to-day, public municipal council}
D14	municipal school	{ISCED97, children & young, ISCED1+2, formal teaching, municipality}
D16	0-9th form	{ISCED97, children & young, ISCED1+2, compulsory}
D17	0th form	{ISCED97, children & young, ISCED1+2, compulsory, preparation}

D18	1-9th form	{ISCED97, children & young, ISCED1+2, compulsory, general basic}
D19	10th form	{ISCED97, children & young, ISCED1+2, optional}

For applying the plurality of cognitive models which according to the present may be used, the following systematic operation may be performed in advance:

1. All feature values existing in two country-specific ontologies to be aligned, are respectively registered in a country-specific matrix.
2. If feature values in the two matrices to be aligned are completely overlapping (e.g. "ISCED0-pre-primary" in DK and "ISCED0-pre-primary" in GE in Tables 1 and 2), the feature columns in question should be merged into one column.
3. If a feature is possessed by a concept, the numeric value should be "1", otherwise "0" in the matrices.
4. If a feature value in one matrix is completely included in a feature value in the other matrix (e.g. "ISCED1+2" in DK and "ISCED1" in GE), a concept possessing the feature that includes the other feature (e.g. Danish "ISCED1+2") should have numeric value "1" in both feature columns (e.g. "ISCED1+2" in DK and "ISCED1" in GE). It means that a concept possessing a feature value that is included in the other feature (e.g. German "ISCED1") should have numeric value "1" only in the feature column in question.
5. If feature values in the Danish and German matrices are partly overlapping (e.g. "ISCED1+2" in DK and "ISCED2+3" in GE), a dummy column referring to the exact overlapping feature value (e.g. "ISCED2" for both DK and GE) is created. In this example, a Danish concept possessing a feature "ISCED 1+2" should have numeric value "1" in both "ISCED 1+2" and "ISCED2" columns, but not in the "ISCED2+3" column.

In this way, the country-specific matrices in this case representing the Danish, and German educational systems are respectively generated from the TOs and the algorithms may be applied hereto to infer the relations between CSCs

Datacreation from non-TO

In the following the Japanese and Danish educational systems are taken as an example. However, the dataset belonging to a first and second culture can be extracted
5 from various sources as described earlier and of course to any subject or culture set of interest

The datasets consist of educational terms and their definitional features that are manually extracted from text corpora. The Japanese corpora used for this experiment
10 are: **1)** "Outline of the Japanese School System" published by the Center for Research on International Cooperation in Educational Development (CRICED), University of Tsukuba, Japan; and **2)** "Higher Education in Japan" published by the Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT). The Danish documents are downloaded from the Euridice web-site published by the Education,
15 Audiovisual and Culture Executive Agency under the EU commission. These corpora are written in English and hence it is feasible to identify original expressions of educational terms in the respective languages from existing parallel- or content aligned corpora. This enables one to eventually achieve translation between Japanese CSCs and Danish CSCs through the English term mapping.

20 The CSCs and their definitions, all written in English, are manually extracted from the text corpora, e.g. the Danish CSC "municipal school (DA: *folkeskole*)" and its definition "*a comprehensive school covering both primary and lower secondary education, i.e. one year of pre-school class, the first (grade 1 to 6) and second (grade 7-9/10) stage basic education, or in other words it caters for the 6-16/17-year-olds*". From this
25 definition, a feature list consisting of "comprehensive school" "primary and lower secondary education" "basic education" "targeted for 6-16/17 years old" is created. This definition also implies that "municipal school" is categorized into three sub-CSCs "preschool class", "first stage" and "second stage", respectively having their features "one year preschool education" "1-6 grades" and "7-9/10 grades". These sub-CSCs are
30 supposed to inherit features defined in the superordinate CSC, in this case "municipal school". In this way, 59 Danish CSCs and 54 Japanese CSCs and their features all written in English are listed up. In addition, some features are manually standardized, e.g. a feature "continuing education for adults" in Denmark is standardized with a feature "opportunities for life-long learning" in Japan. Finally, in order to handle the

problem of numeric feature values which the system does not inherently handle, the same principles for creating the matrices defined above.

Accordingly, in total 229 features are registered in the two matrices, respectively representing the Danish- and Japanese educational systems. In each matrix, if a feature is possessed by a CSC, the numeric value "1" appears, otherwise "0" is assigned. In these matrices, the Danish- and Japanese CSCs are respectively denoted as D_i and J_j and feature IDs are assigned as f_k . Both the Danish- and Japanese CSCs and their features are alphabetically registered.

Fig. 3a illustrates some embodiments of the present method, i.e. the Cross-Categorization Approach (CCA) for aligning semi-structured or unstructured domain knowledge existing in two cultures such as two heterogeneous legal systems (i.e. the educational systems belonging to legal districts, Japan and Denmark). In some embodiments, Jaccard similarity coefficient and the BMG are employed, for computing the degree of relations between preferably all possible combinations of concepts existing in the two legal systems (cultures). In other words, fig. 3a depicts an example of a workflow of the present invention for structuring semi-structured or unstructured data sources existing two heterogeneous domain knowledge, while analyzing inter-relation between the respective data sources.

The alignment component in fig. 3a accommodates algorithms for inferring relations between CSCs either in a symmetric manner or an asymmetric manner. When among others the BMG is employed as the inference algorithm, the inference algorithm computes probabilities that an information receiver (concept learner) possessing one of the domain knowledge successfully infers the meaning of a CSC belonging to the other domain knowledge. When among others the Jaccard similarity coefficient is employed as the inference algorithm, relations are inferred in a symmetric manner.

Fig. 3b depicts a flowchart illustrating an example wherein the present method is applied to the automatic learning of an ontology from a semi- or unstructured data source belonging to one culture. This can be achieved by applying the inference algorithm to the same data source to compute similarity scores among concepts within the same data source.

The structuring component in both Figs. 3a and 3b accommodate algorithms for identifying latent ontological (hierarchical) structures existing in the two independent semi- or unstructured data sources, while inferring relations between the CSCs in the two independent data sources.

5

To the similarity (probability) scores obtained from the inference algorithms, is in an advantageous example applied an extended version of the IRM, the so-called normal Infinite Relational Model (n-IRM) proposed by (Herlau et al., 2012), in order to cross-categorize the educational concepts existing in the two legal systems. The application of the n-IRM achieves to identify number of categories, i.e. groups of educational concepts for the respective legal systems, and to compute the degree of the relations between the categories in the two legal systems. Finally, the IRM is applied to the original data consisting of legal concepts and their features (inferential links), in order to identify underlying relationships, i.e. structures of inferential links, behind the specified concept system, i.e. the identified categories.

10
15

The flowchart of the described approach depicted in Fig.3a illustrates that the present approach simultaneously categorizes e.g. legal concepts existing in two legal systems and from there to structure two independent concept systems that are inter-operable in the most efficient manner. By employing the generalization model, e.g. as exemplified the BMG, it is illustrated that the present approach provides the inferential mechanism of legal concept mapping where a reasoner B (Japanese) generalizes and interprets a new legal concept A (Danish concept) introduced from the legal district A (Denmark) by comparing the new concept A and his or her background knowledge of legal concepts B belonging to the legal system B (Japanese legal system), and vice versa.

20
25

Example extraction of features and CSC

Exemplary Datasets used can be obtained from e.g. UIS who collected data from UNESCO Member States on an individual basis as seen in fig. 4. The purpose of collecting data, according to UIS is to *map the Member States' national education systems according to the International Classification of Education (ISCED)*. UIS aims for *Member States to report their data in an internationally comparative framework*. These datasets from all over the world are downloadable from UIS' web-site¹ One of the challenges of using these datasets is how to map the numeric feature values of

30
35

dimensions such as “starting age” and “cumulative duration of education.” For example, in the Danish educational system, the starting age of upper secondary school is defined as “16-17 years old” and its cumulative years of education is “12-13 years”. On the other hand, the Japanese educational system is a so called “single-track system” meaning that the starting age of upper secondary school is exactly defined as “15 years old” and its cumulative years of education is “12 years”. To handle this difficulty in an objective and systematic manner, the following procedure has been implemented:

1) If a feature value in one country is completely included in a feature value in the other country (e.g. a feature “6-12 y.o.” in Japan is completely included in a feature “6-17 y.o.” in Denmark), a term possessing the feature that includes the other feature (a term possessing “6-17 y.o.” should also possess “6-12 y.o.”), and **2)** If two features from the respective countries are partly overlapping (e.g. “13-15 y.o.” in Japan and “14-17 y.o.” in Denmark), a dummy feature referring to the exact overlapping range (i.e. “14-15 y.o.”) is created. In this example, a Japanese term that possesses “13-15 y.o.” should also possess the dummy feature “14-15 y.o.” In the same . Here, Japanese and Danish datasets have been used for the analysis. Each dataset consists of educational terms defined by several pre-defined feature dimensions such as ISCED level, programme destination and orientation, starting age, cumulative duration of education, and entrance requirements. Most feature dimension values are pre-defined, i.e. for the programme destination dimension, values are pre-defined as [general | prevocational | vocational]. 1

<http://www.uis.unesco.org/education/ISCEDmappings/Pages/default.aspxway>, a Danish term that possesses “14-17 y.o.” should also possess the dummy feature “14-15 y.o.”. In order to objectively assess feature-based similarity measures, simpler datasets that do not contain these ambiguous feature dimensions/values have been prepared as control data. It means that these simpler datasets only contain the standardized feature dimensions/values defined by UIS. Based on these, similarity scores are computed by applying algorithms as discussed in above.

Example: Implementation of the cross-categorization approach

Fig. 5 depicts an overview of the results obtained from the n-IRM applied to similarity scores computed by Jaccard similarity coefficient and the BMG. The four plots (1-a; 1-b; 1-c; 1-d) in the upper row illustrate similarity scores computed in all combinations of concepts between the Japanese- and the Danish educational systems, while the four

plots (1-e; 1-f; 1-g; 1-h) in the bottom row shows the cross-categorization results obtained by the n-IRM computation.

The four plots (2-a; 2-b; 2-c; 2-d) in fig. 6 and the four plots (3-a; 3-b; 3-c; 3-d) in Fig 7 contrast mean values and standard deviation of each cluster in the four plots (1-e; 1-f; 1-g; 1-h) in fig. 5 obtained by the n-IRM. The four columns from the left to the right in Fig 5 illustrate the results respectively representing the followings:

1. First column (1-a; 1-e; 2-a; 3-a)): similarity scores computed by Jaccard Similarity coefficient when the Japanese educational concepts are set as reasoner's background knowledge;
 2. Second column (1-b; 1-f; 2-b; 3-b): similarity scores computed by Jaccard similarity coefficient when the Danish educational concepts are set as reasoner's background knowledge;
 3. Third column (1-c; 1-g; 2-c; 3-c): similarity scores computed by the BMG when the Japanese educational concepts are set as reasoner's background knowledge; and
 4. Fourth column (1-d; 1-h; 2-d; 3-d): similarity scores computed by the BMG when the Danish educational concepts are set as reasoner's background knowledge;
- The plots (1-a and 1-b) depicting the results obtained from Jaccard similarity coefficient are identical, since Jaccard similarity coefficient equally consider features possessed by legal concepts in the two educational systems. The cross categorization results of the Jaccard scores in 1-e and 1-f illustrate that the patterns of co-clusters (intersections between Danish clusters and Japanese clusters) identified in 1-e and 1-f are almost identical, although the number of clusters obtained in 1-e and 1-f are slightly different. The slight differences of the numbers are likely caused during the cluster assignment process employing the Chinese Restaurant Process. For example, such phenomenon can be seen in an intersection between the first Japanese cluster and the second Danish cluster in 1-e and in an intersection between the second Japanese cluster and the second Danish cluster in 1-f. The difference in the intersections in 1-e

and in 1-f is that the Japanese cluster in 1-e is split up into two clusters in 1-f, which generated the difference in the number of clusters obtained in 1-e and in 1-f. Despite the differences in the cluster-numbers, the patterns of the co-clusters are almost identical. Accordingly, the distribution of grey scaled means and standard deviations between 2-a and 2-b as well as 3-a and 3-b are relatively similar in fig. 6 and fig. 7.

On the other hand, the results obtained from the BMG shown in the plots 1- c and 1- d are substantially different, since the BMG only considers and weights features of concepts existing in reasoner's background knowledge. These differences in similarity scores clearly influence on the cross-categorization results obtained by the n-IRM applied in the present example as a relational model. The number of clusters obtained in 1-g is 13 for the Japanese system and 12 for the Danish system, while 8 for the Japanese and 9 for the Danish in 1-h. The obtained numbers of clusters are correlative to the number of concepts that are considered as reasoner's background knowledge.

To be more specific, when the Japanese educational system consisting of 54 educational concepts is considered as reasoner's background knowledge in 1-g, the number of the obtained clusters are larger compared to the situation where the Danish educational system consisting of 27 educational concepts is considered as reasoner's background knowledge in 1-h. The differences in the obtained number of clusters in 1-g and 1-h may for example be caused by the distributions of similarity scores. When the 54 Japanese concepts are considered as reasoner's background knowledge, the features possessed only by these 54 concepts are computed. Hence, differentiations in the similarity scores are stronger across the 54 Japanese concepts rather than across the 27 Danish concepts.

This phenomenon is identified in 1-c where the horizontal lines are more visible compare to the 1-d. Accordingly, the partition of the Japanese concepts results in the ne-grained clusters, which also affect the partition of the Danish concepts in 1-g during the generative process of the n-IRM. In the same way, when the Danish educational system is considered as reasoner's background knowledge, differentiations in the similarity scores are stronger across the 27 Danish concepts, which can be seen in the stronger vertical lines in 1-h. Thus the partition of the Danish domain knowledge consisting of only 27 concepts affects the partition of the 54 Japanese concepts in 1-h, which results in fewer clusters.

The mean values in Fig 6 show that the clusters obtained from the n-IRM. The grey scale indicates that, when a cluster is close to black, the mean value is close to one, and vice versa. The plots 2-a and 2-b showing the results from the Jaccard scores are more uniform, i.e. the majority of clusters are grey coloured. On the other hand, in the plots 2-c and 2-d, the gray colors are more differentiated due to the weight assigned to each feature during the computation of the conditional probabilities in the BMG. Hence, the interactivity between Japanese- and Danish clusters is clearly explained with the results obtained from the BMG, while it is more ambiguous with the results obtained from the Jaccard scores.

Finally, the standard deviations shown in fig. 7 explain the uniformity within each cluster. If a cluster is completely uniform, a cluster in each plot indicates with the white colour, and vice versa. Fig. 7 does not indicate substantial differences in the grey colour distributions among the four plots (3-a; 3-b; 3-c; 3-d). An implication identified here is that, when the clusters are fine grained, the proportion of the light grey coloured clusters are slightly dominant as is indicated in 3-c. This also may imply that the fine-grained uniform clusters obtained from the BMG are potentially more effective for interactively uncovering latent hierarchical structures of respective domain knowledge.

Alignment/relation and structuring of the two domain knowledge

Fig 8, fig 11 and fig 14 all illustrate the co-clustering results obtained from the n-IRM, combined with the implementation of the IRM analyses for the original concept-features matrices of the respective legal system (culture) after immobilizing the concept clusters obtained from the n-IRM. The plots at the center of fig 8, fig 11 and fig 14 respectively corresponds to the plots 1-f, 1-g and 1-h in fig 5.

In fig 8, fig 11 and fig 14, the right- and the bottom sides of the n-IRM plots show feature clusters that respectively represent the contents (feature structure) of the Japanese- and the Danish clusters obtained from the n-IRM.

These feature clusters are obtained by applying the IRM to the original binary matrices consisting of two modes, i.e. legal concepts and their features, while immobilizing the Danish- and Japanese concept modes, i.e. the legal concept clusters obtained from the n-IRM respectively applied to the Jaccard and the BMG similarity dataset.

Fig 9, fig 12 and fig 15 show mean values of intersections of legal concept clusters obtained from the n-IRM and eta values (highest likelihoods) of feature clusters obtained from the IRM, respectively corresponding to fig 8, fig 11 and fig 14.

5 fig 10, fig 13 and fig 16 respectively list: i) in the left column, members of the Japanese- and the Danish legal concept clusters and feature clusters with eta values equal or over 0.5 respectively extracted from fig 9, fig 12 and fig15; and ii) in the right column, members of the Japanese- and the Danish feature clusters.

10 These results demonstrate interesting differences among the similarity measures employed in the present example according to the invention. In fig 10 employing the Jaccard similarity coefficient, a Japanese concept cluster J3 including junior college, regular course, junior college, correspondence course, College of technology, regular course does not specifically relate with any Danish concept clusters in Fig 8 and Fig
15 9.

 However in the daily legal practice, a Danish reasoner may have to accredit Japanese students who hold a degree from the J3 education and desire to study in the Danish educational systems. In Fig 13 employing the BMG, two Japanese concept
20 clusters J10 consisting of junior college, regular course, junior college, correspondence course, College of technology, regular course; and J11 consisting of university, undergraduate, university, undergraduate, pharmacy, medicine etc., university, undergraduate, correspondence course, respectively have stronger relation with D12 consisting of short cycle tertiary education, and D7 consisting of medium cycle tertiary
25 education and Bachelor's program. These clusters are obtained from a Japanese reasoner's viewpoint, i.e. the background knowledge of the Japanese educational system is taken into consideration. When inspecting the feature clusters in fig13, the feature clusters differentiating the Japanese concept clusters J10 and J11 are respectively consisting of: Jf3, whose features are among others programme
30 destination B and ISCED 5, short; and Jf10, whose feature is programme destination A.

 Hence the Danish concept cluster formations are also affected by these feature dimensions. More specifically, the Danish concept clusters D12 and D7 are distinguished by the fact whether the feature cluster Df9: programme destination B is
35 possessed or not. On the other hand, in fig16 employing the BMG, J7 consisting of

junior college, regular course, junior college, correspondence course, college of technology, regular course also includes university, undergraduate and university, undergraduate, correspondence course. This is because the Japanese J7 concept cluster has formed as a result of alignment based on a Danish reasoner's background knowledge where feature possessed only by the Danish educational concepts have been taken into consideration. The J7 concept cluster is aligned to the Danish concept cluster D4 with the mean value 0.29 in fig15. Fig14 shows that the intersection of J7 and D4 is uneven, i.e. higher standard deviation. This implies that these two concept clusters are rather linked based on individual concept-concept relations.

Elaborating the issue of higher standard deviations, some of the intersections between the Japanese- and the Danish concept clusters indicate lower mean values with higher standard deviations as shown in e.g. the intersection between J1 and D1 in plot 2-d in fig 6 and plot 3-d in fig 7. It means that this intersection consists of similarity scores that are substantially uneven influenced by the combinations of members in J1 and D1. For example, fig 17 shows the similarity relations between the members of the legal concept clusters J1 (column) and D1 (row) obtained from the n-IRM applied to the BMG similarity scores when Danish legal concepts are considered as reasoner's background knowledge. The members of both J1 and D1 consists of different types of concept such as primary school, Master's degree and Doctor's degree concepts. This indicates that, by analyzing the individual similarity relations between members of J1 and D1, our approach enables us to identify more fine-grained relationships within the two clusters.

Fig 18 and fig 19 are concept systems, the so-called Terminological Ontologies (TOs), developed based on the principles and methods proposed by Madsen et al. (2004a). The principles of the Terminological Ontology (TO) by Madsen et al. (2004a) defines several rules for developing ontologies. Most importantly, a category must always inherit features possessed by its superordinate concepts. This approach is fairly intuitive and reasonably consistent with the hierarchical structure of categories that are generally discussed by cognitive scientists such as Murphy (2004). Madsen et al. (2004a) also defines that, when a category is divided into several sub-categories, these subcategories must be differentiated by one or more feature(s) possessed by each sub-category. This also implies that the category and a sub-category must also be differentiated by one or more specific feature possessed by the sub-category in

question. Another important principle is that the TO approach allows polyhierarchy structures so that one sub-category may be related to two or more superordinate categories. The principles of TO also defines more strict rules derived from the traditional view of terminology that aims at achieving strict standardization. For example, the principle of uniqueness of dimension defines that a given dimension for dividing a category into several sub-categories may only occur once in an ontology. Madsen et al. (2004) argues that this uniqueness of dimensions further strict rules based on logical approach for the purpose of standardization, we employed some of these few important rules that enable us to visually represent the extracted feature structures from our method in a systematic way.

Fig 18 represents the Japanese legal concept system developed upon the Japanese feature structure where the eta values equal or above 0.5 are extracted in fig 12, while fig 19 is developed for representing the Danish legal concepts system when the Japanese legal concepts are aligned with the Danish legal concepts based on the Danish feature structure in fig 15. The ontologies contain all the concept clusters (categories) obtained from the n-IRM computation, which are displayed in the grey boxes. The corresponding concept clusters in the other legal system, which are identified with our cross-categorization approach (i.e. mean values coloured in fig 12 and fig 15), are listed above the concept clusters in question. For example in fig 18, the Japanese concept cluster (category) J11 consisting of members referring undergraduate educations corresponds to the Danish concept cluster D7 consisting of Danish education concepts referring to Bachelor level educations in Denmark. Some of the grey boxes in the ontologies are listed as e.g. category share Jf11 in fig 18. More specifically, this category consists of the four sub-categories: J2, J9, J10, J11 (right upper part of the table) that share the feature cluster Jf11 in fig 12. This means that one of the superordinate categories of J11: undergraduate education cluster is the superordinate category *category share Jf11*, that are divided into the four sub-categories: J2: Master level educations, J9: sort-term tertiary education, J10: vocational tertiary education and J11: undergraduate general tertiary education. White boxes in fig 18 and fig 19 describe feature structures of each concept category. For example, the concept category *category share Jf11* can be represented as the group of concepts that commonly possess features related to ISCED5 (refers to general university education). On the contrary, the concept category J11: undergraduate education has a polyhierarchical structure which inherits features from other two

superordinate categories, category share Jf10 possessing features related to destination A (refers to academic path) and category share Jf4 possessing features related to ISCED4 (refers to vocationally-oriented post upper secondary education). The more specific meanings of the J11: undergraduate education can be described based on all the contents of feature clusters that represent J11 from the list shown in fig 13. Eventually, the semantic differences between J11 and its corresponding Danish concept category D7 can also be described by contrasting all the contents of feature clusters in 11 respectively representing these two concept clusters in question.

As mentioned above, fig 18 represents the Japanese legal concept system based on the Japanese feature structure where the eta values equal or above 0.5 are extracted in fig 12. It means that the Japanese ontology in fig 18 is developed after the Danish legal concepts are aligned with the Japanese legal concepts based on the Japanese reasoner's background knowledge. In the same way, the terminological ontology in fig 19 is developed for representing the Danish legal concepts system after the Japanese legal concepts are aligned with the Danish legal concepts based on the Danish reasoner's background knowledge in fig 15. Contrasting these two ontologies, another interesting aspect is uncovered. In Section9, the number of concept clusters obtained from the different similarity scores, i.e. similarity scores computed by i) Jaccard similarity; ii) similarity scores computed by the BMG when the Japanese educational concepts are set as reasoner's background knowledge; and iii) similarity scores computed by the BMG when the Danish educational concepts are set as reasoner's background knowledge. In case of ii), the obtained Japanese concept clusters were fine-grained since the number of the Japanese educational concepts registered in the dataset is large so that categorization criteria within the Japanese educational system become specific. In other words, in order to distinguish 54 educational concepts, several specific features are used to differentiate each educational concept. Since the BMG only consider features possessed by the Japanese educational concepts in the case of ii), the cross-categorization dimensions (rather strictly following the ISCED levels) are highly influenced by the Japanese way of categorizing the educational concepts, which resulted in the fine-grained cluster partitions both for the Japanese- and the Danish concepts.

Accordingly, the Danish concept categories obtained from the n-IRM are aligned with specific Japanese concept categories in fig 18. On the other hand, in case of iii), the

BMG is applied to dataset consisting of fewer concepts including the adult education concepts. When observing fig 19, the Japanese concept cluster J3: upper secondary evening courses is mapped with the four Danish concept clusters (categories): D2: upper secondary, general education; D3: upper secondary, vocational education; D5: adult education; and D8: practical admittance course for 5B. The hierarchical structure in Fig 19 implies that the superordinate concept of D2 and D3; D2 and D5; and D5 and D8 are respectively: category share Df7 (refers to ISCED 3, i.e. upper secondary education); category share Df9 and Df10 (refer to destination A, general education, i.e. academic general education); and category share Df6 and Df11 (refer to starting age over 18 y.o.). In other words, the Japanese concept category J3 could be mapped with the Danish categories from these different dimensional views when the concept category is viewed by a Danish reasoner. Although this structure can also be observable from fig 15, the terminological ontology in fig 19 enables us to visually investigate such semantic relationships.

Theory of COM, SL and TL

First of all, the COM framework involves the at least two parties; one being the communicator who is conveying meanings of an source language (SL) concept to an information receiver who is receiving a target language (TL) stimulus that is supposed to be a translation of the SL concept in question. The COM framework may deal with domain-specific knowledge that is culturally-rooted in a specific country, e.g. the educational system, social system, legal system, traditional events etc. For convenience, the COM framework assumes that the average population in a specific country has general knowledge e.g. about the educational system in his/her country, as domain knowledge. Such country-specific knowledge is in most cases officially translated into English. Hence the English expression of each country-specific concept which also possesses an original local expression is considered as input data in this scenario. By identifying a country- and domain-specific corpus officially written in English, it is possible to manually extract English expressions of concepts and their definitions. These English expressions and their features that are identified in their definitions can be used for constructing a hierarchical structure of categories based on the basic principles described in Murphy (2004). These basic principles of forming a hierarchical structure of categories are, as a starting point, assumed to be consistent with the principles of Terminological Ontology whose methodological basis is laid out in the next chapter. Although some principles of Terminological Ontology may interfere

with the natural hierarchical formation of categories, the assumption here is that the Terminological Ontology might still be a useful method to be employed in the framework. Terminological ontologies are constructed both for the SL- and TL domain knowledge, respectively considered as the communicator- and the information receiver's prior knowledge.

Contrasting to the Relevance Theory of Communication, such prior knowledge is considered as *context C*. We can draw two types of scenarios where a communicator, based on his/her prior knowledge, is going to identify an appropriate translation from new objects existing in a TL information receiver's cultural domain; and where an information receiver, based on his/her prior knowledge, is going to generalize the meanings of original concepts from stimuli given by a communicator. The former could be considered as the SL-oriented communication and the latter as the TL-oriented communication. It means that, e.g. in case of the TL-oriented communication, if the information receiver has his/her prior knowledge about the educational system in his/her country, this knowledge is considered as *context C*. The information receiver is supposed to have no knowledge about the educational system in the SL culture. The SL communicator is now providing a stimulus that is a TL translation of an SL educational concept. This TL translation appears as *a new encountered information P* to the information receiver's *context C*. The union of *P* and *C* is supposed to generate the contextual effect according to the Relevance Theory of Communication. In other words, the union of *P* and *C* implies the information receiver's assumption *Q* about the new information *P*, which is the communicator's intention. Here, if a cognitive environment is shared by two people, the set of all facts is manifest to both communicator and audience and therefore this may possibly generate a common ground based on the symmetric coordination. However, in a realistic scenario, the two parties use different languages and are mastering different concepts so that the way people construct mental representations and perform inference are inherently different. Thus, and as stated previously, it is most realistic and easiest to achieve the asymmetric coordination. In order for the TL information receiver to infer and generalize the original meaning of the SL concept, the category-based inductions, i.e. feature-based similarity measures, are applied as algorithms. For example, the model of computing similarities based on features proposed by Tversky (1977) enables one to compute such asymmetric similarities. To re-emphasize, this asymmetric similarity algorithm explains the views of Rips (1975) and Osherson et al. (1990) that induction

from X to Y is not in general the same as from Y to X . Thus, again, the similarity of a given category to a target category is uni-directional.

The above is somehow forming the “ideal picture”, of a cognitive framework consisting of at least four elements required for the COM framework. These four key elements are: **a)** the asymmetric co-ordination; **b)** the contextual effects generated based on the union of a new object P and prior knowledge C ; **c)** the taxonomic organization of categories; and **d)** the category-based induction. These elements are integrated into the COM framework as shown in Figures 20c and 20d. These figures respectively depict the SL-oriented communication and the TL-oriented communication described above. Figures 20a and 20b, respectively illustrate how the two hierarchically structured concept systems are mapped depending on the communication patterns. To fulfill the requirements for the elements **a)** and **b)**, methods for representing element **c)**: the taxonomic organization of categories, should be identified. In addition, algorithms for performing element **d)**: category-based induction should be identified.

Asymmetric cross-cultural communication

In Figure 21a, the left- and the right columns respectively represent the asymmetric cross-cultural communication illustrated in Figures 21b and 21c. For example, the left-upper graph shows that a Japanese communicator who has prior knowledge of the Japanese educational system considers that “D48: single structure education”, “D19: first stage”, “D36: municipal school” and “D44: private school” are the most similar concepts to the Japanese elementary school. However, from the viewpoint of a Danish audience who has prior knowledge of the Danish educational system, D48 (Danish compulsory education consisting of primary and lower secondary levels) and D19 (the first part of the single structure corresponding to the primary education, however, this concept is not so common as the single structure system in Denmark) have higher relevance to the Japanese elementary school. Fascinatingly, the Japanese communicator in Figure 21a identifies “D12: continuation school (DA: *efterskole*)”, “D21: youth school – full-time system” as the most similar concepts to the Japanese lower secondary school. In Denmark, the concept of “lower secondary school” does not exist, because the lower secondary level is included in the single structure education. The concepts which the Japanese identified Glückstad F.K (2012) *The 26th Annual Conference of the Japanese Society for Artificial Intelligence, 2012 - 9* - are alternative educations targeted for young people in the age bracket of 14-17 years old. Thus, if the Japanese communicator selects “continuation school (DA: *efterskole*)” as translation

for conveying meanings of the Japanese lower secondary school, the Danish audience might imagine other meanings than the ones the Japanese intended to convey.

Contrary, the right-lower graph shows that “D48: single structure” is the most relevant concept to the Japanese lower secondary school from the viewpoint of the Danish

5 audience. In this way, the cognitive simulation could potentially identify a translation candidate from an audience’s viewpoint. Such a feedback function might be useful for, e.g. a pivot translation system employed for Machine Translation (MT) and Cross-Lingual Information Retrieval (CLIR).

10 The diagrams in fig 21a and other similar representations of relations between CSCs are related to the matrix of fig. 2c. In the matrix of fig. 2c the calculated relation between CSCs are given by a value $a - i$. In Fig. 21a the relation is given by the height of the bar.

15 **More about TO and application to COM:**

A TO is developed within a knowledge sharing community, then dynamically updated and validated. If it is necessary to share knowledge with other communities, TOs developed in different communities should be compared, aligned and merged as needed. Based on this view, terminological ontologists argue that concepts are defined
20 in a language dependent context, and therefore TOs are inherently language dependent.

The principles of the TO have been developed in the research and development project called CAOS - Computer-Aided Ontology Structuring – where the aim has been to develop a computer system designed to enable semi-automatic construction of
25 ontologies (Madsen, Thomsen and Vikner, 2004; 2005). The uniqueness of the TO is its feature specifications and subdivision criteria (Madsen et al., 2004a; 2004b). The use of feature specifications is subject to principles and constraints described in detail in (Madsen et al., 2004a). Most importantly, a concept automatically inherits all feature specifications of its superordinate concepts. According to Madsen et al. (2004a), *this principle models the principle of traditional terminology that ‘the intension of the subordinate concept includes the intension of the superordinate concept’* (ISO 704: 5.4.2.2; cf. also Madsen 1999: 21). Secondly, subdivision criteria that have been used for many years in terminology work are strictly implemented in the TO by introducing dimensions and dimension specifications (Madsen et al., 2004a; 2004b). This enables
30 the CAOS prototype to perform consistency checking which helps in constructing TOs.
35

A dimension of a concept is an attribute occurring in a non-inherited feature specification of one or more of its subordinate concepts. Values of the dimension allow a distinction among sub-concepts of the concept in question. For example, a dimension of the concept “academic degree” is [LENGTH OF EDUCATION] whose values are [2-3 year | minimum 4 years]. These dimension values distinguish the sub-concepts: “junior college” and “university”. The dimension can only occur in feature specifications on sister concepts and a given value can only appear on one of these sister concepts. This second principle implies the third principle, that is, a concept must be distinguished from each of its nearest superordinate concepts by at least one feature specification. In the TO, a concept dimension and its feature values are registered as (DIMENSION : [value1, value2, ...]). In the case of Fig. 5 *dimension specification* of a concept “academic degree” is represented as (LENGTH OF EDUCATION : [2-3 year | minimum 4 years]). This dimension specification subdivides the concept “academic degree” into two sub concepts “junior college” and “university” which respectively possess primary features, [LENGTH OF EDUCATION : 2-3 year] and [LENGTH OF EDUCATION : minimum 4 years]. The features that subdivide these two concepts are called *primary feature specifications* which are differentiated from other feature specifications that are inherited from superordinate concepts. It is also allowed to define one or more dimension specification of a concept, e.g. an example shown as the concept “degree” in Fig. 22. In this way, a concept must be distinguished from each of its nearest superordinate concepts as well as from each of its sister concepts by at least one feature specification (Madsen et al., 2004a; 2004b).

These principles are fairly intuitive and reasonably consistent with the hierarchical structure of categories described by (Murphy, 2004). On the other hand, the principles of the TO also defines more strict rules derived from the traditional view of terminology that aims at proper standardization. For example, a principle of uniqueness of dimension defines that *a given dimension may only occur on one concept in an ontology*. Madsen et al. (2004a) argues that *uniqueness of dimensions contributes to create coherence and simplicity in the ontological structure, because concepts that are characterized by means of a certain common dimension must appear as descendants of the same superordinate concept*. In the same way, Madsen et al. (2004) also defines the uniqueness of primary feature specifications as *a given primary feature specification can only appear on one of the daughters*. The argument is that these uniqueness principles *make it possible to a certain extent to carry out automatic placing*

of concepts into an ontology. Another point is that the TO principles allows polyhierarchy structures so that one concept may be related to two or more superordinate concepts. This principle is inconsistent with the principle of the taxonomic organization of categories described in (Murphy, 2004).

5 Although some of the principles are consistent with the principles of human taxonomic organization of categories described in (Murphy, 2004), the question arises whether these strict artificial rules are applicable to the COM framework when the framework aims at simulating cognitive processes of human category-based inductions. In a way, this issue is quite similar to Temmerman's (2000) question about
10 "fuzziness of categories" arguing that categories cannot be absolutely classified by logical and ontological means.

 The explanation of elements of the present invention can be initiated from the viewpoint that Terminological Ontology (TO), the method that has been introduced by
15 (Madsen et al., 2004) at CBS, contributes to identify an optimally relevant translation by assisting one to systematically organize conceptual features from domain knowledge (corresponding to the taxonomic organizations). Such systematically organized features can be used for linking an SL concept with a TL concept based on a plurality of cognitive models as algorithms of aligning two culturally-dependent taxonomies.

20 TO is a domain-specific ontology used for knowledge sharing, which normally is applied in terminology work within the domain of language for special purposes. The unique points of TO that differentiate it from other types of ontologies are its feature specifications and subdivision criteria. A feature specification consists of a feature dimension and its value. Thus, a representation of a whole concept is a feature
25 structure, i.e. a set of feature specifications corresponding to the unique set of characteristics that constitutes that particular concept. Terminological ontologists argue that concepts are defined in a language-dependent context, and therefore, TO is language- or culturally dependent. TO is developed within a knowledge sharing community, then dynamically updated and validated. If it is necessary to share
30 knowledge with other communities, TOs developed in different communities should be compared, aligned and merged as needed.

Application of IRM prior to BMG

The first strategy (fig 23a) is in a way the most natural approach to judge how an ontology is learned from data consisting of CSCs and features that respectively represent specific domain knowledge existing in two cultures. Thus the IRM is directly applied to the CSC-feature matrices, respectively created from the aforementioned English corpora describing the Danish- and the Japanese educational systems. Accordingly, 59 Danish CSCs and 229 features are simultaneously clustered into 5 and 10 categorical classes. In the same way, 54 Japanese CSCs and 229 features are respectively clustered into 6 and 11 categorical classes.

By this application of IRM some categorical classes (e.g. Danish classes 3, 4 and 5; and Japanese classes 1, 3, 5, and 6) are successfully formed only with CSCs that are related to the respective categorical classes such as “upper secondary”, “open education”, “secondary”, and “lower secondary”. However, the rest of the categorical classes are partly formed with CSCs that represent different categorical classes Glückstad F.K (2012) *The 26th Annual Conference of the Japanese Society for Artificial Intelligence, 2012*. For example, the Danish categorical class 1 consists of CSCs that are supposed to belong to “pre primary” and “adult education” and Japanese categorical class 2 consists of CSCs that are supposed to belong to “tertiary” and “primary”.

The successful Danish categorical class 3 “upper secondary” has a very dense relationship with feature cluster 7 consisting of “16-18 years old” and “post compulsory education” and with feature cluster 10 consisting of “upper secondary education” and “vocational perspectives”. In the same way; the Danish categorical class 5 representing degree programs targeted for adults has a dense relationship with feature cluster 6 consisting of features: “opportunities for lifelong learning”, “part time”, “possibilities for combining education and work”, “occupational function”, and “open education”. The resulting data also shows another notable point that the Japanese categorical classes 1 and 3 both have a dense relationship with feature cluster 9 consisting of “non-compulsory educational school” and “post-compulsory education”. However, the Japanese categorical class 1 - “upper secondary” - has also a strong relationship with feature cluster 8 consisting of “16-18 years old”. Also the Japanese categorical class 3 - “alternative post compulsory” - has another relationship with feature cluster 10 consisting of “education + practical training”. This indirectly indicates that the Japanese categorical classes 1 and 3 both belong to a super-ordinate category (although it does

not exist in the dataset) referring to “post-compulsory education”. This kind of information could be useful for representing knowledge in a taxonomical structure, e.g. for constructing Terminological Ontologies [Madsen 2004]. The results of experimental strategy 1 indicate that if few decisive features exist for representing a categorical class, the IRM effectively sorts CSCs that relate to these decisive features. However, when relationships between categorical classes and feature clusters are weak, there is a tendency that CSCs that belongs to different categorical classes start to be mixed into one class. However, this combination of IRM followed by BMG makes it possible to analyze how features and each categorical class are related to each other.

Application of BMG prior to IRM

A second strategy (Fig 23b) is to apply the BMG to directly compute similarity relations between CSCs existing in the two cultures, and thereafter to apply the IRM in order to cluster CSCs in the respective countries into categorical classes. This enables us not only to observe the inter-relations of categorical classes existing in the two cultures, but also to instantly scrutinize more specific similarity relations between each category member (i.e. CSCs) existing in the two cultures. Accordingly, 59 Danish CSCs and 54 Japanese CSCs are simultaneously clustered into 11 and 9 categorical classes, respectively (fig. 24a). Figure 24a shows the graph sorted according to extracted assignments of categorical classes computed by the IRM algorithm.

The results in fig. 24b show that both the Danish- and the Japanese CSCs are clustered into a more fine-grained level compared with the results obtained from the first experimental strategy. Almost all members in each categorical class are grouped together with other members that are supposed to belong to the same categorical class. For example, some CSCs such as the Japanese “J3: college of technology (JP: *koto-senmon-gakko*)” and the Danish “D36: municipal school (DK: *folkeskole*)” are CSCs that are difficult to be categorized in a multi-cultural context. While, in the first- and second experimental strategies, these CSCs have been included in a more ambiguous larger categorical class where CSCs that are supposed to belong to different categorical classes have been grouped together, J3 and D36 are respectively grouped into a more specific and independent categorical class, i.e. the Japanese categorical class 6 and the Danish categorical class 4, in this third strategy. When observing the η -sorted data it is possible to study more complex relationships of

categorical classes in a cross-cultural context. For instance, the Japanese categorical class 6 where “J3: college of technology” belongs, has a strong relationship with the Danish categorical class 2 “upper secondary” class, but also has a little weaker relationship with both the Danish categorical classes 7 and 8, which respectively represent “vocational academy” and “vocational college” categories providing a 2 years post-secondary degree in Denmark. The creation of an η -sorted data may further provide a clear picture of how each country-specific categorical class is related to categorical classes existing in another country in a very complex and comprehensible manner. This kind of overview of how categorical classes in different cultures are inter-related is highly useful and valuable not only for mapping CSCs but also for constructing ontologies in a multi-cultural context.

One advantage of the second strategy is that the direct application of the BMG enables us to analyze further specific similarity relations between category members of the respective categorical classes existing in the two cultures. Figure 24c illustrates how the category members of the Japanese categorical class 6 in Figure 24b are related with the category members of the Danish categorical classes 2, 7 and 8. As discussed in the previous section, the η -sorted data (not shown) shows that the Japanese categorical class 6 has the strongest relationship with the Danish categorical class 2 and slightly weaker relationship with the Danish categorical classes 7 and 8. Figure 24c explains these relationships between the classes by showing that all the category members of the Danish categorical class 2 share at least one feature with all the category members of the Japanese categorical class 6, while only 75% of the category members of the Danish categorical classes 7 and 8, respectively, share at least one feature with 75% of the category members of the Japanese categorical class 6. On the other hand, when observing individual relationships between category members between the Japanese- and the Danish categorical classes, similarity relationships are not necessarily strong in most of the combinations. Here, the selection of feature-based similarity measures plays in to the considerations.

In this example BMG is applied for feature-based similarity measure. However, for implementing the IRM based on the second strategy, it is possible to apply other feature-based similarity measures, such as the Jaccard similarity coefficient [Tan, 2005], [Jaccard, 1901] and Tversky’s set-theoretic model [Tversky, 1977], which

compute similarities based on common features and distinctive features possessed by the two CSCs in question.

5 The present work implies, for example, the features that influence the formation of categorical classes in both cultures could be prioritized as necessary feature dimensions when constructing TOs. This may prevent eliminating important features that could be used for computing similarities based on the BMG. An attractive aspect of the IRM is that the model can be a more complex clustering of three or more relations simultaneously. Hence, this can be applied for multi-cultural modelling as well. These
10 results imply that the integration of all methods, i.e. the BMG+IRM+TO approach, can enable us not only to map CSCs by respecting nuances of each concept existing in each respective culture, but also to construct TOs that are cross-culturally interoperable as well as mono-culturally clarified.

15 Fig 23c is an elaborated version of fig. 3a and shows two different approaches of mapping CSCs. From the view-point of the CSC mapping in cross-cultural communication, the BMG+IRM which is a cognitive model of generalization is designed to compute such fuzziness of the human mind by reflecting prior knowledge of a learner who compares a new object with something he/she knows in advance. On the
20 other hand, an advantage of the TO+BMG approach is the clarification of domain knowledge in a specific knowledge community and the ability to compare the *hierarchical structures* across different communities. The BMG combined with the IRM, both of which originate from the cognitive sciences, is designed to compute the fuzziness of human mind by reflecting prior knowledge of a learner who compares a
25 new object with something he/she knows in advance. By considering the advantage of the TO that is the clarification of domain knowledge in a mono-cultural knowledge community, the BMG+IRM+TO approach may be an optimal solution, which may enable us not only to map CSCs by respecting nuances of each concept in the respective cultures, but also to construct TOs that are multiculturally interoperable.

30

Terminological approach and COM in relation to existing theories

The terminological approach considers a concept consisting of several feature specifications, which possesses one or several lexical expressions, as starting point.
35 The approach is indeed reasonably consistent with the way concepts are represented

by features and with the basic principles for forming a hierarchical structure of categories described in Murphy (2004). On the other hand, Sperber & Wilson (1986) considers an assumption as a set of structured concepts and emphasizes logical functions of concepts. Thus they distinguish logical entry and encyclopaedic entry, which are the information attached to a concept. More specifically, Sperber & Wilson argue that encyclopaedic entries typically vary across speakers and they are open-ended. On the other hand, *logical entries are small, finite and relatively constant across speakers and times..... the content of an assumption is constrained by the logical entries of the concepts it contains, while the context in which it is processed is, at least in part, determined by their encyclopaedic entries* (Sperber & Wilson, 1986: 88-90). This implies that the encyclopaedic entries in a way correspond to the concept representation in Murphy (2004). However, the way the context is processed is highly influenced by the lexical semantic view in Sperber & Wilson (1986). The general lexical semantics considers first a lexical item that possesses several meanings as starting point. The view Sperber & Wilson take is rooted in the lexical semantic approach and is implied from their considerations that *concepts have both logical and lexical entries that provide a point of contact between input and central processes, between the linguistic input system and the deductive rules of the central conceptual system. Recovery of the content of an utterance involves the ability to identify the individual words it contains, to recover the associated concepts, and to apply the deductive rules attached to their logical entries. ... We assume, then, that the "meaning" of a word is provided by the associated concept.... This allows us to maintain a somewhat ecumenical view of lexical semantics* (Sperber & Wilson, 1986:90). This difference in views of the terminological approach by Murphy (2004), and the lexical semantic approach by Sperber & Wilson (1986), results in different inferential approaches, deductions and inductions. Sperber & Wilson's RTC framework employs the *deductive* system based on the elimination rules as inferential algorithm. On the other hand, Murphy (2004) argues that the inferential process involved in communication is a category-based *induction*. For example, for the RTC framework, the elimination rule for generating a conclusion from a premise can be illustrated in the following way:

Mother-elimination rule

Premise: (X-mother-Y)

Conclusion: (X-female parent-Y)

Here, it is considered that *the meaning of a word is provided by a definition which expresses the individually necessary and jointly sufficient conditions for the word to apply. For instance, the definition of "mother" could be "female parent". If this is so, it can be represented by assigning "mother" as the lexical entry for the concept "female parent"* the elimination rule in the above logical expression (Sperber & Wilson, 1986: 90). On the contrary, Murphy (2004: 241) argues that *one of the major uses of categories is to make predictions about novel items*. He emphasizes that speakers can assume that listeners will, by referring to something with a simple category name, retrieve information about the category and use it to comprehend the meaning. The reason the term *induction* is used for referring to this process is that it involves drawing an uncertain inference of the listeners to the category as a whole. Especially, people are attempting to draw conclusions about one category or a new encountered item based on their prior knowledge of another category. In this inductive process, people use similarity judgment. According to Heit and Rubinstein (1994: 420), *prior knowledge could be used dynamically to focus on certain features when similarity is evaluated. In this conception, inductive reasoning is an active process in which people identify features in the premise and the conclusion categories that are relevant to the property being inferred*. This statement indicates that, if prior knowledge is organized in a hierarchical structure based on the terminological approach, such prior knowledge could effectively be used for the inductive reasoning as pointed out by Heit and Rubinstein.

To summarize, while Figure 25a illustrates the original RTC framework, Figure 8b proposes a revised version of the RTC framework that forms the basis of the COM framework.

Hence, now it is possible to draw a more concrete picture of the COM framework that is an integration of the selected elements of the Relevance Theory of Communication and the replaced elements of the Knowledge Effects. First of all, the COM framework involves the two parties; one being the communicator who is conveying meanings of an SL concept to an information receiver who is receiving a TL stimulus that is supposed to be a translation of the SL concept in question. The COM framework only deals with domain-specific knowledge that is culturally-rooted in a specific country, e.g. the educational system, social system, legal system, traditional events etc. For convenience, the COM framework assumes that the average population in a specific country has general knowledge e.g. about the educational system in his/her country, as domain knowledge. Such country-specific knowledge is in most cases officially

translated into English. Hence the English expression of each country-specific concept which also possesses an original local expression is considered as input data in this scenario. By identifying a country- and domain-specific corpus officially written in English, it is possible to manually extract English expressions of concepts and their definitions. These English expressions and their features that are identified in their definitions can be used for constructing a hierarchical structure of categories based on the basic principles described in Murphy (2004). These basic principles of forming a hierarchical structure of categories are, as a starting point, assumed to be consistent with the principles of Terminological Ontology whose methodological basis is laid out in the next chapter. Although some principles of Terminological Ontology may interfere with the natural hierarchical formation of categories, the assumption here is that the Terminological Ontology might still be a useful method to be employed in the framework. Terminological ontologies are constructed both for the SL- and TL domain knowledge, respectively considered as the communicator- and the information receiver's prior knowledge.

Contrasting to the Relevance Theory of Communication, such prior knowledge is considered as *context C*. We can draw two types of scenarios where a communicator, based on his/her prior knowledge, is going to identify an appropriate translation from new objects existing in a TL information receiver's cultural domain; and where an information receiver, based on his/her prior knowledge, is going to generalize the meanings of original concepts from stimuli given by a communicator. The former could be considered as the SL-oriented communication and the latter as the TL-oriented communication. It means that, e.g. in case of the TL-oriented communication, if the information receiver has his/her prior knowledge about the educational system in his/her country, this knowledge is considered as *context C*. The information receiver is supposed to have no knowledge about the educational system in the SL culture. The SL communicator is now providing a stimulus that is a TL translation of an SL educational concept. This TL translation appears as *a new encountered information P* to the information receiver's *context C*. The union of *P* and *C* is supposed to generate the contextual effect according to the Relevance Theory of Communication. In other words, the union of *P* and *C* implies the information receiver's assumption *Q* about the new information *P*, which is the communicator's intention. Here, if a cognitive environment is shared by two people, the set of all facts is manifest to both communicator and audience and therefore this may possibly generate a common ground based on the symmetric coordination. However, in a realistic scenario, the two

parties use different languages and are mastering different concepts so that the way people construct mental representations and perform inference are inherently different. Thus, and as stated previously, it is most realistic and easiest to achieve the asymmetric coordination. In order for the TL information receiver to infer and generalize the original meaning of the SL concept, the category-based inductions, i.e. feature-based similarity measures, are applied as algorithms. For example, the model of computing similarities based on features proposed by Tversky (1977) enables one to compute such asymmetric similarities. To re-emphasize, this asymmetric similarity algorithm explains the views of Rips (1975) and Osherson et al. (1990) that induction from X to Y is not in general the same as from Y to X . Thus, again, the similarity of a given category to a target category is uni-directional.

The above is somehow forming the “ideal picture” of a cognitive framework consisting of at least four elements required for the COM framework. These four key elements are: a) the asymmetric co-ordination; b) the contextual effects generated based on the union of a new object P and prior knowledge C ; c) the taxonomic organization of categories; and d) the category-based induction. These elements are integrated into the COM framework as shown in Figures 8a and 8b. These figures respectively depict the SL-oriented communication and the TL-oriented communication described above. Figures 20a and 20b, respectively illustrate how the two hierarchically structured concept systems are mapped depending on the communication patterns.

The multi-dimensional view Murphy (2004), may naturally support the IRM approach (Kemp et al., 2006). Murphy (2004) argues that when several categorical features are related, the cluster of these features form a so-called family resemblance structure. *This can happen either through the relations of prior knowledge (Ahn, 1990; Kaplan and Murphy, 1999; Spalding and Murphy, 1996) or through induction that relate the features (Lassaline and Murphy, 1996). Murphy and Allopenna (1994) shows that when the features of a category formed a consistent set, the category was much easier to learn than when they were inconsistent or simply neutral.* Their findings in a way indirectly explain how the IRM played a role in the above IRM+MBG and MBG+IRM approaches **1)** applying the IRM directly to two CSC-feature matrices respectively representing educational domain knowledge in Japan and Denmark for first categorizing them into categorical classes that are to be afterwards compared and aligned; **2)** applying the BMG to directly compute similarity relations between CSCs in

the two cultures, thereafter applying the IRM for clustering CSCs in the respective cultures into categorical classes.

The empirical results from the above strategy 1) could be interpreted based on Murphy and Allopenna's work, that, if a consistent set of features is formed among CSCs, a category has been successfully learned. On the other hand, if feature sets for CSCs were inconsistently structured, e.g. few relevant- and many irrelevant features, the clustering results have not been optimal. The question is then how to interpret the results obtained from the above strategy 2). It seems that the following suggestion by Murphy and Allopenna is key for answering this question. More specifically, they suggest that *knowledge helps learning because it relates the features in the category, rather than through the properties of the features themselves* (Murphy, 2004: 151). In addition, Kaplan and Murphy's (2000) work shows that when thematic features are present in categories, background knowledge is helpful even though the knowledge is incomplete or imperfect. Subjects are also able to ignore features that are inconsistent and still be able to use the most accurate knowledge. The second strategy first applies the BMG. This implies that the loosely structured sets of features representing culturally-specific domain knowledge, either of the Japanese or the Danish educational system, is considered as prior knowledge, and the other part is considered as new information compared against the prior knowledge. This process identifies all existing links between the Danish CSCs and the Japanese CSCs when they share at least one common feature. Thus, if a Japanese who has knowledge about a categorical class consisting of Japanese CSCs that provide "compulsory education", the IRM likely categorizes Danish CSCs that also provide "compulsory education" and creates a link at the categorical class level. Such links for co-clustering categorical classes in the two cultures are found in η sorted data.

To summarize, the initial framework of COM described in this application is assumed to consist of the following modules: **1)** identification of domain specific content-aligned (or parallel) corpora, preferably consisting of for example SL-English and English-TL language combinations; **2)** terms and features extractions from content-aligned corpora which describe domain specific terms and their definitions for the respective domain-knowledge in the SL- and TL cultures; **3)** construction of ontologies for the respective domain-knowledge in the SL- and TL countries based on the information extracted in 2); **4)** creation of feature structures for each concept for the respective domain-knowledge in the two countries, and standardization of feature labels used in the two

culture-specific ontologies; **5)** alignment of the two structured feature sets based on feature-based ontology mapping algorithms applying cognitive models, i.e. Tversky's contrast model (Tversky, 1977) and eventually the Bayesian Model of Generalization (Tenenbaum & Griffiths, 2001); and **6)** identification of corresponding translation

5 candidates from the content-aligned corpora consisting of SL-English and English-TL language combinations. For convenience, the initial COM framework is schematically illustrated in Figure 26.

10

15

Items

- 5 1. A Method for inferring relations between cultural specific concepts (CSC) in two cultures at least comprising the steps of
- extracting and listing said cultural specific concepts (CSCs) and features of said CSCs from at least a first corpora belonging to a first culture and a second corpora belonging to a second culture,
- 10 - applying a algorithm to infer relations between said CSCs in the first and the second corpora.
2. The method according to item 1 wherein a pivot language is chosen and where at least the cultural specific concepts and features in the at least first and second corpora
- 15 are translated to the pivot language.
3. The method according to any of the preceding items wherein the algorithm is a Bayesian inference model.
- 20 4. The method according to any of the preceding items wherein the said algorithm is based on connectionist approach such as Artificial Neural Networks.
5. The method according to item 3 wherein the bias is on background knowledge and wherein the significance of a feature is reversely proportional with the occurrence of
- 25 said feature and/or wherein the Bayesian algorithm is applied from the first and/or second culture's point of view.
6. The method according to any of the preceding items further comprising the step of identifying at least one candidate corresponding pair of a CSC from the first culture to
- 30 the second culture.
7. The method according to any of the preceding items further comprising the step of identifying at least one probability that an information receiver belonging to the second corpora successfully infers the meaning of a CSC belonging to the first corpora
- 35 translated to the second corpora.

8. The method according to any of the preceding items further comprising the step of applying an unsupervised algorithm for learning taxonomies, and for structuring said hierarchical relations among said taxonomies and their features.

5

9. The method according to any of the preceding items wherein the said unsupervised learning algorithm is applied prior to or after application of the mapping algorithm.

10

10. The method according to any of the preceding items wherein the cultural specific concepts (CSCs) and features of the first culture are extracted from a first ontology and the cultural specific concepts (CSCs) and features of the second culture are extracted from a second ontology, preferably the ontology is a terminological ontology (TO).

15

11. The method according to item 1 – 10 wherein first inference algorithm is applied to unstructured data sets of cultural specific concepts and features to align the CSCs of the first and second cultures, where after the unsupervised learning algorithm is applied to structure the results and wherein the result from the combined application of the said inference algorithm and the said unsupervised learning algorithm is used to construct at least one ontology.

20

12. The method according to item 1 -11, wherein said method is a computer implemented method, and where said method is being executed by a system comprising one or more devices.

25

13. A computer program for executing a method for inferring relations between cultural specific concepts (CSC) in two cultures at least comprising the steps of

- extracting and listing said cultural specific concepts (CSCs) and features of said CSCs from at least a first corpora belonging to a first culture and a second corpora belonging to a second culture,
- applying a algorithm to infer relations between said CSCs in the first and the second corpora

30

14. The computer program according to item 13, wherein the method is the method according to item 1 – 12.

35

15. Computer program according to item 13 or 14, comprising one or more program modules which alone or together execute one or more of the steps in the method according to the present invention and/or wherein the computer program comprises of
5 locally stored and/or remotely stored and/or at least partly of internet and/or cloud based program modules.

16. A system comprising one or more devices, wherein at least one of said devices comprises software for enabling inferring relations between CSCs from a first and a
10 second culture, wherein said relations preferably is inferred by the method according to item 1-12.

15

20

25

Citations:

- 5 Ahn, W. (1990). Effects of Background Knowledge on Family Resemblance Sorting. *Proceedings of the Twelfth Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum. 149-156.
- Bush, R. R., & Mosteller, F. (1951). A model for stimulus generalization and discrimination. *Psychological Review*, 1951, Vol., 58(6; 6), 413-423.
- 10 Cimiano, Philipp, Montiel-Ponsoda, Elena, Buitelaar, Paul, Espinoza, Mauricio, & Gómez-Pérez Asunción. (2010). A Note on Ontology Localization. In *Journal of Applied Ontology* Vol. 5, No. 2, IOS Press, 127-137.
- Chomsky, N. (1986). *Language and problems of knowledge: The Mnagua lectures*. Cambridge, MA: MIT Press.
- 15 Declerck, T., Krieger, H.U., Thomas, S.M., Buitelaar, P., O'Riain, S., Wunner, T., Maguet, G., McCrae, J., Spohr, D. & Montiel-Ponsoda, E. (2010) Ontology-based Multilingual Access to Financial Reports for Sharing Business Knowledge Across Europe. *Internal Financial Control Assessment Applying Multilingual Ontology Framework*, J. Rooz, J. Ivanyos, Eds. Budapest: HVG Press, 67-76.
- 20 Glückstad F.K. (2012). Bridging Remote Cultures: Influence of Cultural Prior-Knowledge in Cross Cultural Communication. *Proceedings of the 26th Annual Conference of the Japanese Society for Artificial Intelligence*, June 2012, Yamaguchi, Japan.
- 25 Glückstad F.K. and Mørup M., Flexible- or Strict Taxonomic Organization?: Impact on Culturally-Specific Knowledge Transfer. In *Proceedings of the 10th International Conference on Terminology and Knowledge Engineering 2012, Madrid, Spain (2012)*
- 30 Hansen T.J., Mørup M. and Hansen L. K. (2011), Non-parametric Co-clustering of Large Scale Sparse Bipartite Networks on the GPU. In *IEEE International Workshop on Machine Learning for Signal Processing (MLSP), IEEE (2011)*

- Haussler, D., Kearns, M. & Schapire, R. E. (1994). Bounds on the Sample Complexity of Bayesian Learning Using Information Theory and the vc Dimension. *Machine Learning*, 14, 83-113.
- 5 Heit, E. (1998). A Bayesian Analysis of Some Forms of Inductive Reasoning. In: M. Oaksford & N. Chater (Eds.), *Rational Models of Cognition*. Oxford: Oxford University Press. 248-274.
- Heit, E. & Rubinstein, J. (1994). Similarity and Property Effects in Inductive Reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 1994, Vol., 20(2; 2), 411-422.
- 10 Herlau T., Mørup, M., Schmidt, M.N. and Hansen, L. K. (2012), Modelling Dense Relational Data. In *IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*, Santander, Spain (2012)
- Jaccard, P. (1901). Étude comparative de la distribution florale dans une portion des Alpes et des Jura. *Bulletin de la Société Vaudoise des Sciences Naturelles* 37: 547–
- 15 579.
- Kaplan, A. S. & Murphy, G. L. (2000). Category Learning with Minimal Prior Knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 2000, Vol., 26(4; 4), 829-846.
- Kemp, C., Tenenbaum, J.B., Griffiths, T.L. Yamada, T. & Ueda, N. (2006). Learning
- 20 Systems of Concepts with an Infinite Relational Model. *The Twenty-First National Conference on Artificial Intelligence*, 2006.
- Lassaline, Mary E & Murphy G. L. (1996). Induction and Category Coherence. *Psychonomic Bulletin & Review*, 3, 95-99.
- Madsen, B.N., Thomsen, H.E. & Vikner, C. (2004a). Principles of a System for
- 25 Terminological Concept Modelling. *Proceedings of the 4th International Conference on Language Resources and Evaluation*. ELRA, 15-19.
- Madsen, B.N., Thomsen, H.E. & Vikner, C. (2004b). Comparison of Principles Applying to Domain Specific Versus General Ontologies. *Proceedings of Ontologies and Lexical Ressources in Distributed Environments 2004*. ELRA, 90-95.

- Madsen, B.N., Thomsen, H.E. & Vikner, C. (2005). Multidimensionality in Terminological Concept Modelling. *Proceedings of the 7th International Conference on Terminology and Knowledge Engineering*, Copenhagen, 161-173.
- Mitchell, T.M. (1997). *Machine Learning*. New York: McGraw Hill.
- 5 Murphy, G. L. (2004). *The Big Book of Concepts*. Cambridge, Massachusetts: The MIT Press.
- Murphy, G. L. & Allopenna, P. D. (1994). The Locus of Knowledge Effects in Concept Learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 1994, Vol., 20(4; 4), 904-919.
- 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)*
- Osherson, D. N., Smith, E. E., Wilkie, O., López, A. & Shafir, E. (1990). Category-Based Induction. *Psychological Review*, 1990, Vol., 97(2; 2), 185-200.
- 15 Pitman J. (2002). Combinatorial Stochastic Processes. *Notes for Saint Flour Summer School*.
- Rips, L. J. (1975). Inductive Judgments about Natural Categories. *Journal of Verbal Learning and Verbal Behavior*, 14. 665-681.
- 20 Rosch, E. & Mervis, C. B. (1975). Family Resemblances: Studies in the Internal Structure of Categories. *Cognitive Psychology*, 1975, Vol.7(4; 4), 573-605.
- Shepard, R. N. (1987). Towards a Universal Law of Generalization for Psychological Science. *Science*, 237, 1317-1323.
- 25 Spalding, T. L. & Murphy, G. L. (1996). Effects of Background Knowledge on Category Construction. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 1996, Vol., 22(2; 2), 525-538.
- Sperber, D. & Wilson, D. (1986). *Relevance: Communication and Cognition*. Oxford: Blackwell.

Tan P.N., Steinbach M. & Kumar V. (2006). *Introduction to Data Mining*. Boston, MA: Pearson Education, Inc.

5 Temmerman, R. (2000). *Towards New Ways of Terminology Description, the Sociocognitive Approach*. Amsterdam, Netherlands: John Benjamins Publishing Company.

Tenenbaum, J. B. & Griffiths, T. L. (2001). Generalization, Similarity, and Bayesian Inference. *Behavioral and Brain Sciences*, 2001, Vol.24(4; 4), 629-640.

Tversky, A. (1977). Features of Similarity. *Psychological Review*, Vol., 84(4; 4), 327-352.

10 Vossen, Piek. (2004). EuroWordNet: A multilingual database of autonomous and language-specific wordnets connected via an interlingual index. *International Journal of Lexicography* 17.2, 161-173.

15 Vossen, Piek., Agirre, Eneko., Calzolari, Nicoletta., Fellbaum, Christiane., Hsieh, Shu-kai., Huang, Chu-Ren., Isahara, Hitoshi., Kanzaki, Kyoko., Marchetti, Andrea., Monachini, Monica., Neri, Federico., Raffaelli, Remo., Rigau, German., Tescon, Maurizio. & VanGent, Joop. (2008). KYOTO: A system for mining, structuring and distributing knowledge across languages and cultures. *Proceedings of the 6th International Conference on Language Resources and Evaluation*, Morocco, 1462-1469.

Claims

1. A Method for inferring relations between cultural specific concepts (CSC) in two cultures at least comprising the steps of :
- extracting and listing said cultural specific concepts (CSCs) and features of said CSCs from at least a first corpus belonging to a first culture and a second corpus belonging to a second culture,
 - applying an algorithm to infer relations between said CSCs in the first and the second corpus where said inference algorithm is applied to unstructured or loosely structured data sets of cultural specific concepts and features to relate the CSCs of the first and second cultures,
 - applying an unsupervised algorithm for learning taxonomies in said CSCs in the first and the second corpus, and for structuring relations among said taxonomies and their features, where said unsupervised learning algorithm is applied to structure the results of the interference algorithm and wherein the result from the combined application of the said inference algorithm and the said unsupervised learning algorithm is used to construct at least one ontology.
2. The method according to claim 1 wherein a pivot language is chosen and where at least the cultural specific concepts and features in the at least first and second corpus are translated to the pivot language.
3. The method according to any of the preceding claims wherein the algorithm is an inference model, such as a Bayesian inference model.
4. The method according to any of the preceding claims wherein the algorithm is based on connectionist approach such as Artificial Neural Networks.
5. The method according to any of the preceding claims wherein the bias is on background knowledge and/or wherein the significance of a feature is reversely proportional with the occurrence of said feature and/or wherein the Bayesian algorithm is applied from the first and/or second culture's point of view.

6. The method according to any of the preceding claims wherein the significance of a feature is biased by background knowledge and/or the occurrence of said features and/or wherein the Bayesian algorithm is applied from the first and/or second culture's point of view.

5

7. The method according to any of the preceding claims wherein the unsupervised algorithm is a relational model.

10

8. The method according to any of the preceding claims further comprising the step of describing differences in meanings between a candidate corresponding pair of a CSC from the first and the second cultures.

9. The method according to any of the preceding claims further comprising the step of visualizing at least one ontology.

15

10. The method according to claim 8 wherein the visualized ontology is constructed by the unsupervised algorithm.

20

11. The method according to any of the preceding claims comprising the step of identifying at least one candidate corresponding pair of a CSC from the first culture to the second culture.

25

12. The method according to any of the preceding claims comprising the step of identifying at least one probability that an information receiver belonging to the second corpora successfully infers the meaning of a CSC belonging to the first corpora translated to the second corpora.

30

13. The method according to any of the preceding claims wherein said unsupervised learning algorithm is applied prior to or after application of the interference algorithm.

14. The method according to any of the preceding claims wherein the cultural specific concepts (CSCs) and features of the first culture are extracted from a first ontology and the cultural specific concepts (CSCs) and features of the second culture are extracted from a second ontology, preferably the ontology is a terminological ontology (TO).

15. The method according to claims 1 -14, wherein said method is a computer implemented method, and where said method is being executed by a system comprising one or more devices.

5 16. A computer program for executing a method for inferring relations between cultural specific concepts (CSC) in two cultures at least comprising the steps of

- extracting and/or listing said cultural specific concepts (CSCs) and features of said CSCs from at least a first corpus belonging to a first culture and a second corpus belonging to a second culture,
- 10 - applying an algorithm to infer relations between said CSCs in the first and the second corpus where said inference algorithm is applied to unstructured or loosely structured data sets of cultural specific concepts and features to align the CSCs of the first and second cultures,
- applying an unsupervised algorithm for learning taxonomies in said CSCs in the first and the second corpus, and for structuring relations among said taxonomies and their features, where said unsupervised learning algorithm is applied to structure the results and wherein the result from the combined application of the said inference algorithm and the said unsupervised learning algorithm is used to construct at least one ontology.

20

17. Computer program according to claim 16 comprising one or more program modules which alone or together execute one or more of the steps in the method.

25 18. Computer program according to claim 16 or 17 comprising locally stored and/or remotely stored program modules and/or consist at least partly of internet and/or cloud based program modules.

19. Computer program according to claim 16, 17 or 18 wherein the method for inferring relations is a method according to claim 1 – 15.

30

20. A system comprising one or more devices comprising a computer program and/or program modules for enabling inferring relations between CSCs from a first and a second culture.

21. System according to claim 20 wherein said devices includes, PCs, tablets and other handheld portable devices, local or remote resources.

5 22. System according to claim 20 or 21 wherein the computer program of the system which infers the relations between CSCs from a first and a second culture is a computer program according to the present invention.

10 23. System according to any of claims 20 - 22 comprising devices for obtaining information from analogue, digital, written and/or audio sources.

24. System according to any of claims 20 - 23 comprising devices for interacting with one or more users receiving, sending and/or processing information.

15

20

1/33

Difficulties in communication

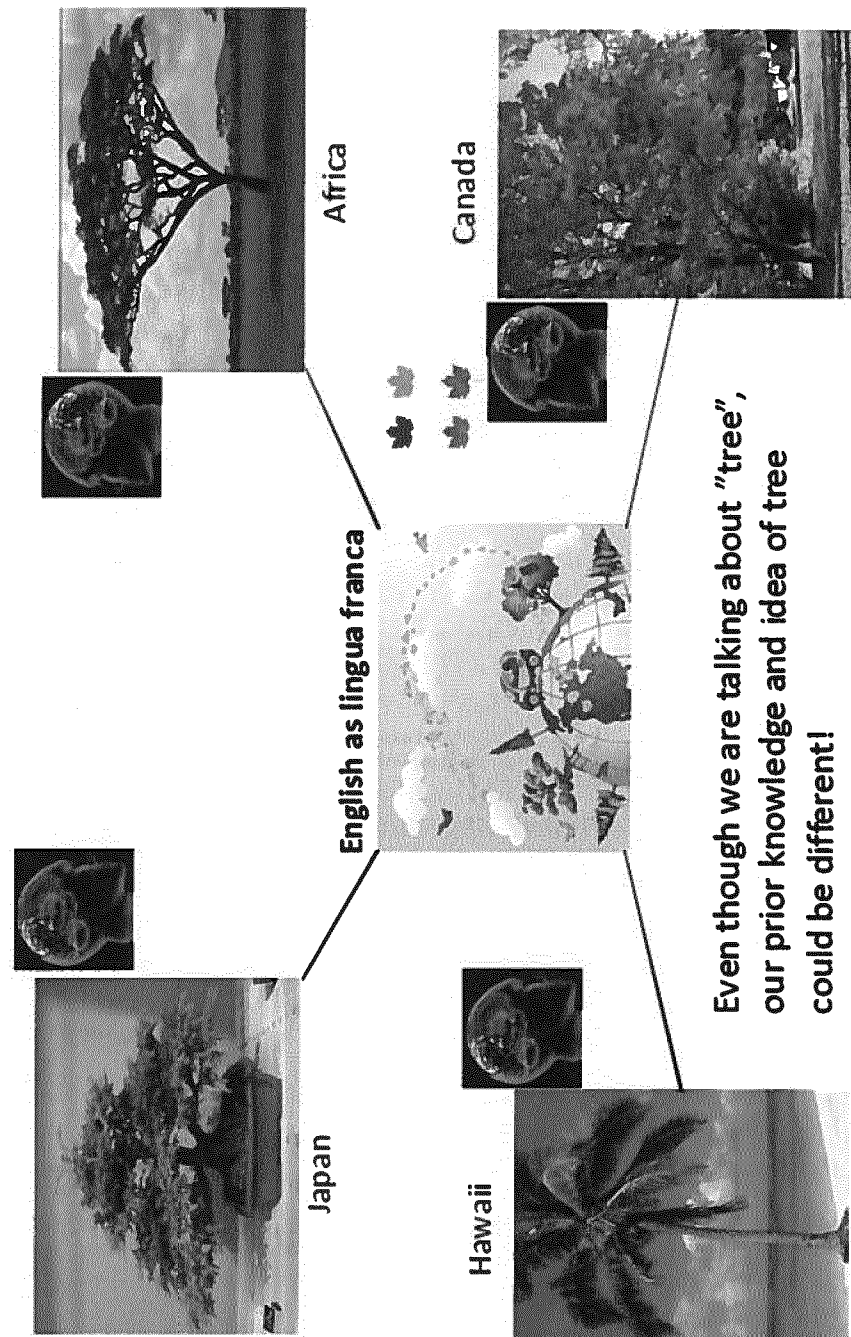


Fig 1

2/33

	Feature 1	Feature 2	Feature 3	Feature 4
J CSC 1	0	1	1	0	
J CSC 2	0	1	0	0	
J CSC 3	1	1	0	1	
....					

Fig 2a

	Feature 1	Feature 2	Feature 3	Feature 4
D CSC 1	0	1	0	0	
D CSC 2	1	1	0	0	
D CSC 3	1	1	0	1	
....					

Fig 2b

	D CSC 1	D CSC 2	D CSC 3
J CSC 1	a	b	c
J CSC 2	d	e	f
J CSC 3	g	h	i

Fig 2c

3/33

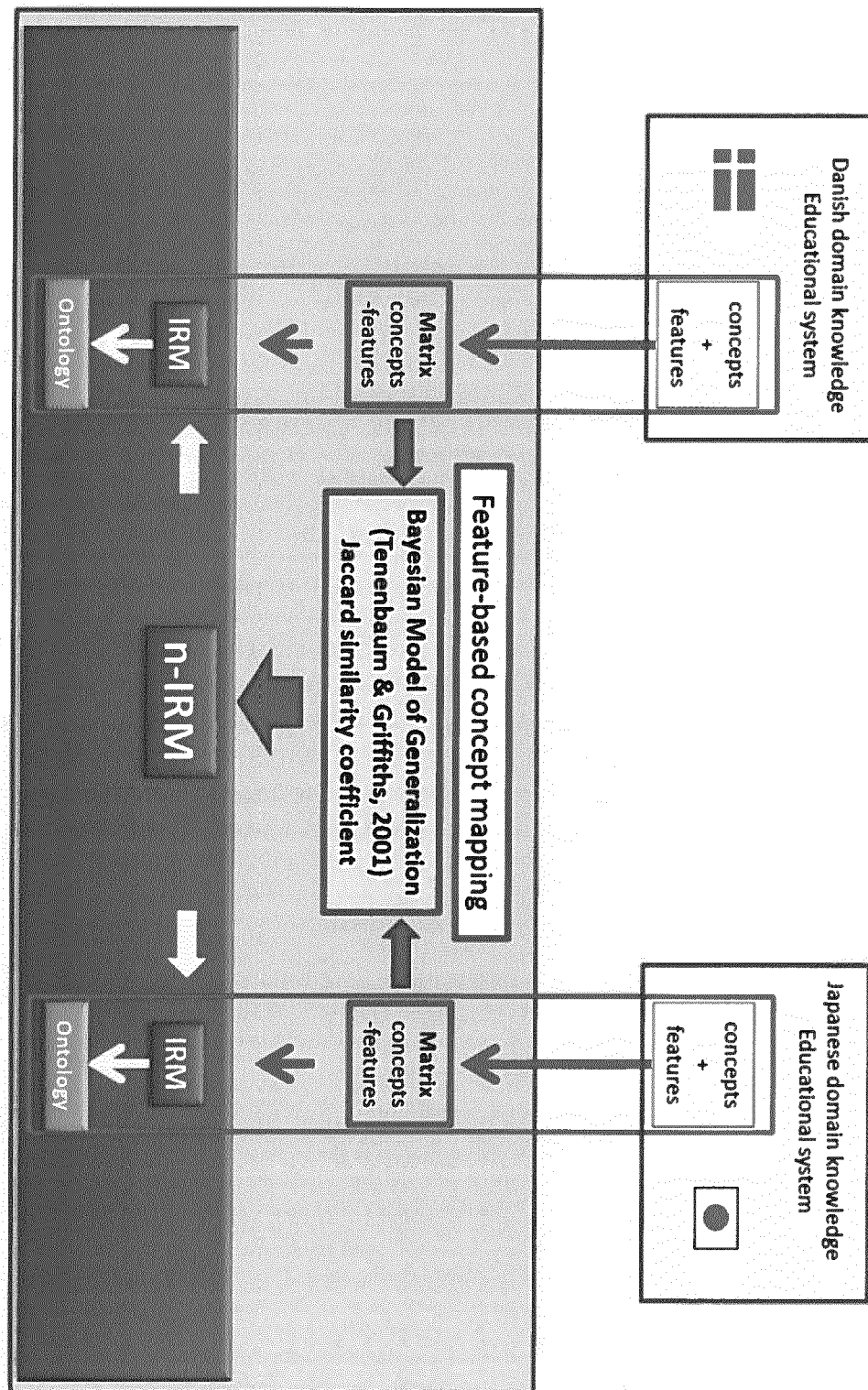


Fig 3a

4/33

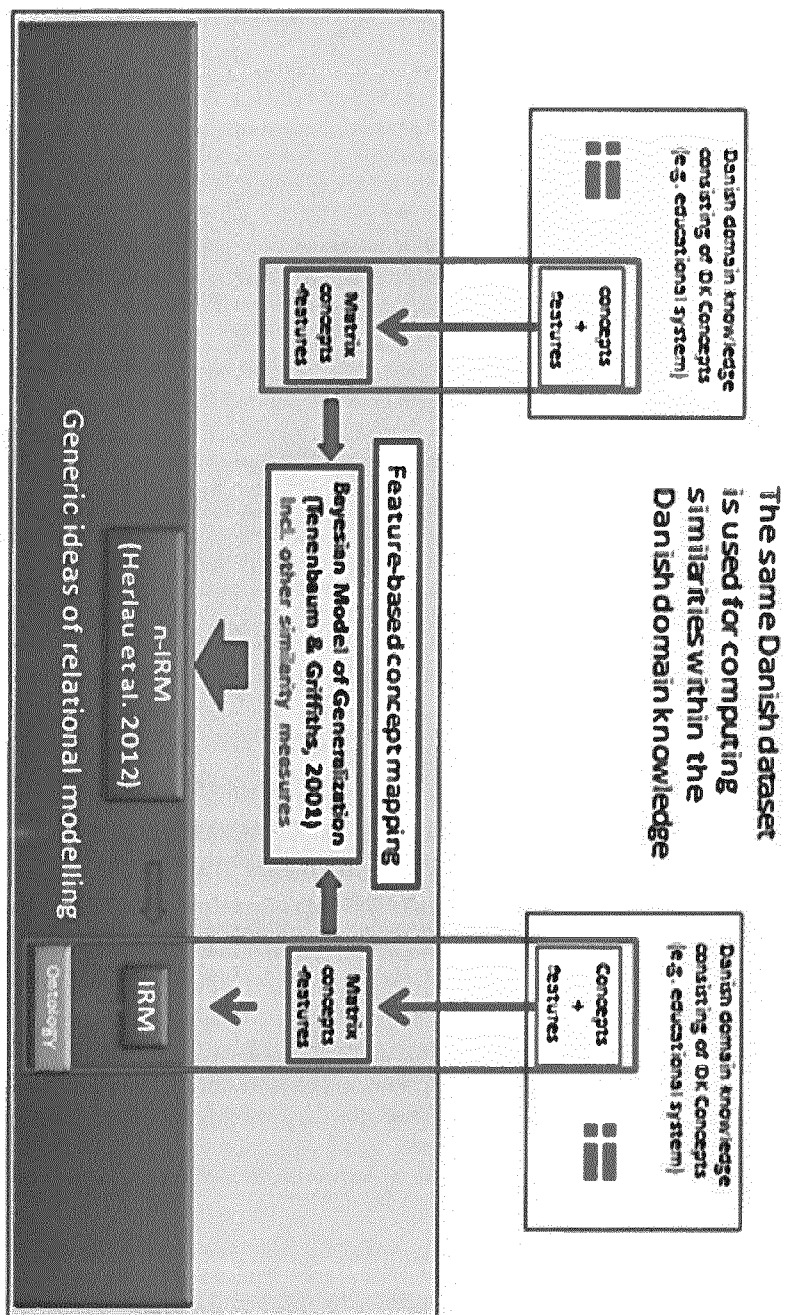


Fig 3b

5/33

Concept ID	ISCED level	Programme destination (A/B/C)	Programme orientation (G/P/V)	Theoretical cumulative duration at ISCED 5	Position in the national degree / qualification structure (Intermediate, first, second, etc...)	Position in the tertiary education structure (Bachelor-Master-PhD)	Minimum entrance requirement (ISCED level or other)	Theoretical starting age	Theoretical duration of the programme	Theoretical cumulative years of education at the end of the programme	Does the programme have a work based element? (Y/N)	Programme specifically designed for adults (Y/N)	Programme specifically designed for part-time attendance (Y/N)
Danish concepts													
D1	0		G					2-5 years	4 years		No	No	No
D2	0		G					5-6 years	1 year		No	No	No
D3	1		G					6-7 years	6 years	6 years	No	No	No
D4	2	A	G				1	12-13	3-4 years	9-10	No	No	No
D7	3	C	V				2A	16-30	3-5 years	14	Yes	No	No
D19	5	B		Short	1st		3A, 3C	18-50	0.5-4 years	13-15	No	Yes	Yes
D20	5	B		Short	1st		3A, 3C	20-30	2-3 years	14	No	No	No
D21	5	A		Medium	1st	Bachelor	3A	18-50	2-4 year	13-15	No	Yes	Yes
D22	5	A		Medium	1st	Bachelor	3A	20-30	3-5 years	16	Yes	No	No
D23	5	A		Medium	1st	Bachelor	3A	20-30	3 years	15-16	No	No	No
Japanese Concepts													
J35	5	B		Short	Intermediate		3 ABC	18	2-3	14-15	No	No	No
J36	5	B		Short Medium	Intermediate		5B	20	1+	15+	No	No	No
J37	5	B		Short	Intermediate		3	18	2-3	14-15	No	No	Yes
J38	5	B		Short	Intermediate		3	18	2	14	No	No	No
J40	5	B					3	18	1+	13+	No	No	No
J41	5	A		Medium	1st	Bachelor	3	18	4	16	No	No	No
J42	5	A		Long	1st	Bachelor	3	18	6	18	No	No	No
J44	5	A		Long	Intermediate		5A 1st, M	22	1+	17+	No	No	No

Fig 4

6/33

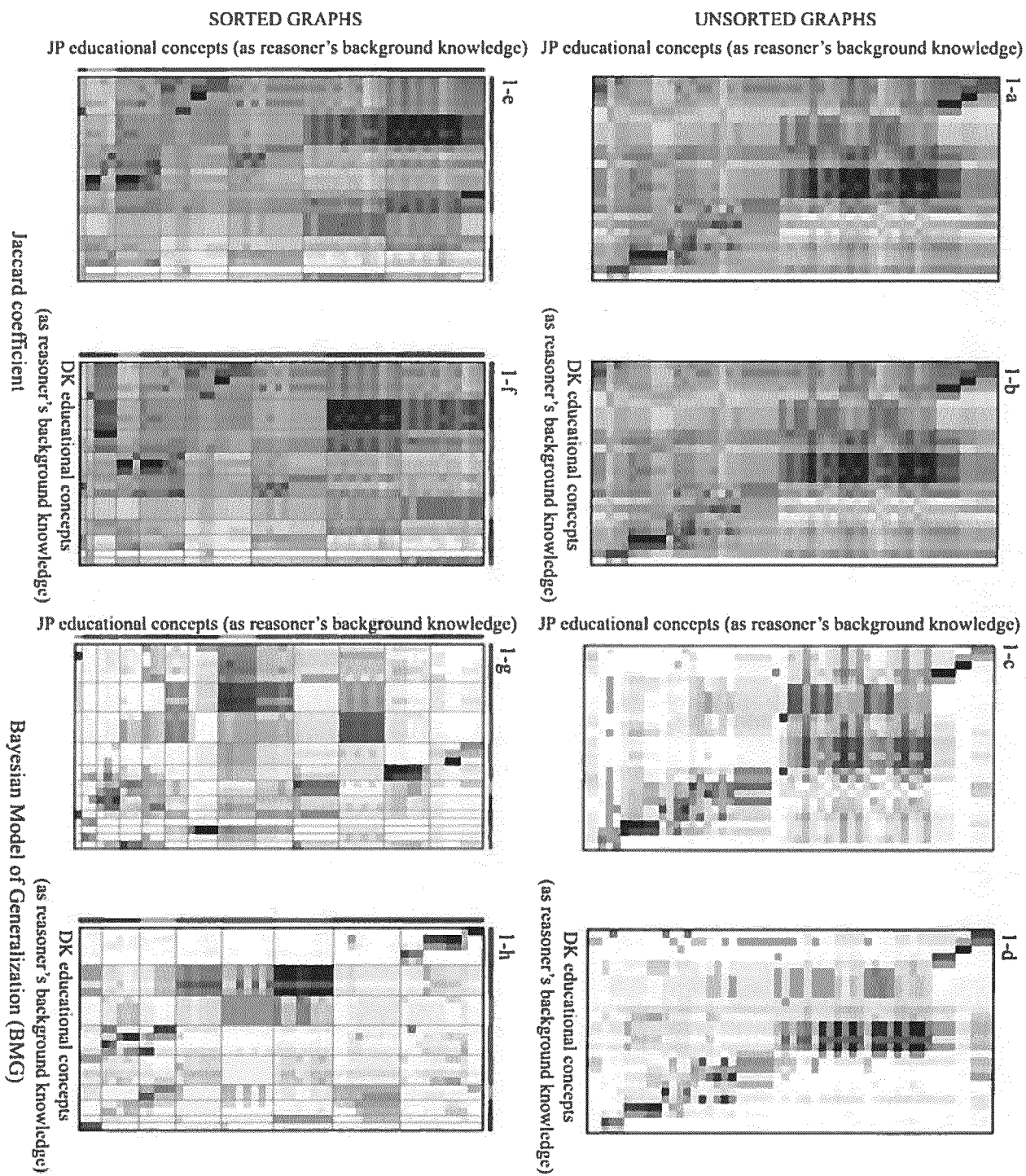


Fig 5

7/33

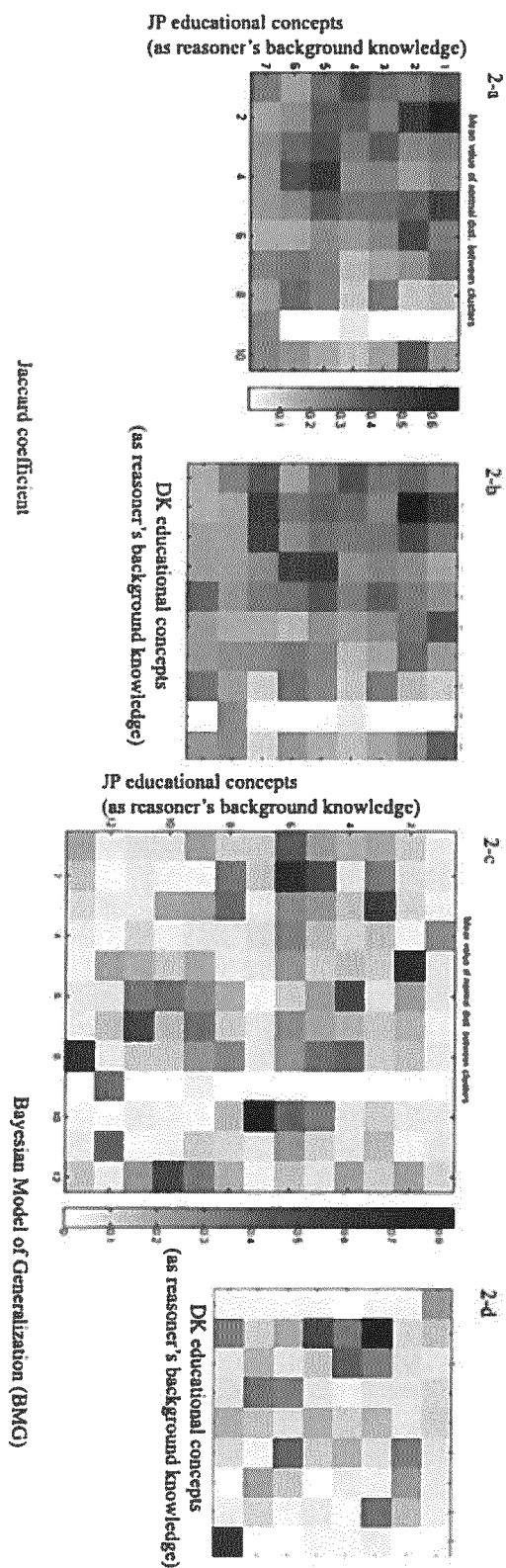


Fig 6

8/33

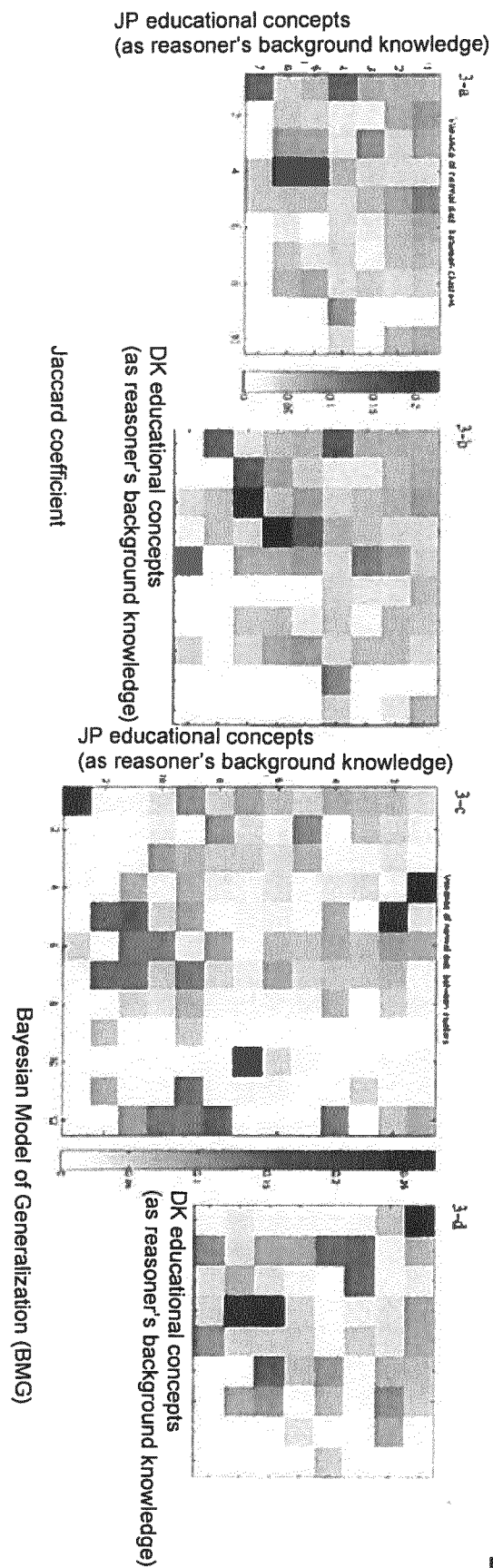


Fig. 7

9/33

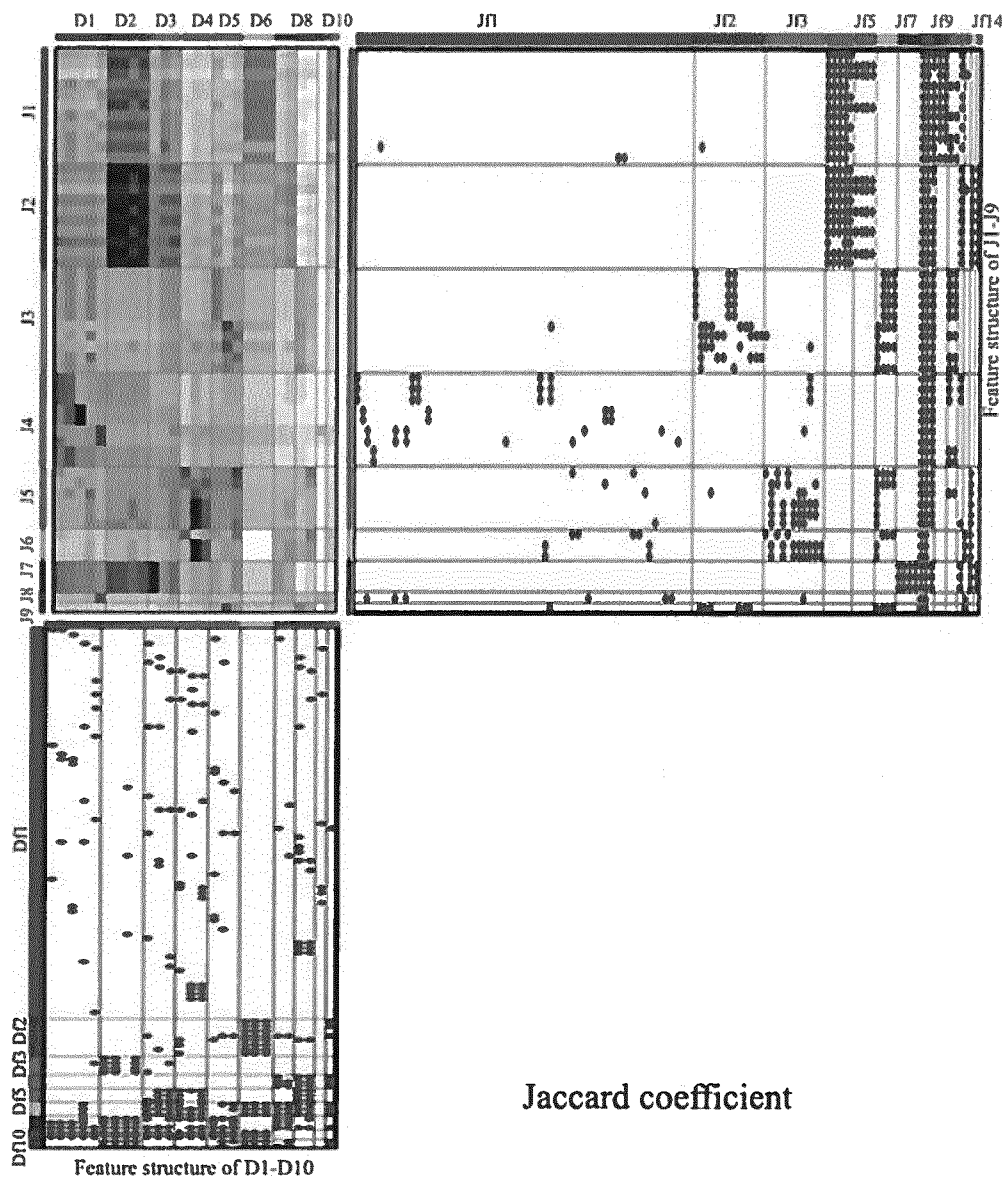


Fig 8

10/33

Avr. 0.25	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	Jf1	Jf2	Jf3	Jf4	Jf5	Jf6	Jf7	Jf8	Jf9	Jf10	Jf11	Jf12	Jf13	Jf14
J1	0.30	0.51	0.34	0.17	0.24	0.44	0.20	0.10	0.00	0.39	0.01	0.01	0.01	0.91	0.27	0.02	0.02	0.94	0.88	0.42	0.62	0.15	0.08	0.08
J2	0.37	0.73	0.46	0.24	0.29	0.31	0.35	0.13	0.00	0.23	0.00	0.01	0.01	0.94	0.48	0.02	0.02	0.94	0.05	0.05	0.92	0.17	0.92	0.92
J3	0.33	0.26	0.31	0.26	0.38	0.24	0.17	0.28	0.00	0.16	0.00	0.32	0.03	0.02	0.02	0.71	0.02	0.97	0.05	0.76	0.08	0.08	0.08	0.08
J4	0.44	0.35	0.30	0.23	0.28	0.19	0.13	0.10	0.04	0.11	0.07	0.01	0.05	0.02	0.03	0.03	0.03	0.96	0.05	0.40	0.30	0.09	0.09	0.09
J5	0.32	0.37	0.32	0.53	0.42	0.22	0.26	0.26	0.00	0.16	0.01	0.03	0.44	0.03	0.04	0.50	0.04	0.95	0.07	0.21	0.25	0.13	0.88	0.13
J6	0.20	0.32	0.26	0.51	0.32	0.13	0.31	0.32	0.00	0.22	0.05	0.02	0.56	0.06	0.07	0.43	0.07	0.60	0.13	0.13	0.20	0.80	0.80	0.20
J7	0.43	0.53	0.51	0.26	0.31	0.16	0.26	0.10	0.00	0.09	0.01	0.02	0.03	0.06	0.07	0.07	0.92	0.91	0.13	0.14	0.80	0.20	0.80	0.80
J8	0.30	0.19	0.21	0.21	0.23	0.16	0.26	0.20	0.25	0.26	0.11	0.07	0.15	0.14	0.17	0.17	0.17	0.60	0.25	0.25	0.33	0.67	0.33	0.33
J9	0.17	0.16	0.21	0.21	0.38	0.21	0.22	0.35	0.00	0.22	0.03	0.47	0.08	0.14	0.17	0.83	0.17	0.60	0.25	0.25	0.33	0.67	0.33	0.33
Df1	0.06	0.01	0.06	0.11	0.07	0.00	0.02	0.11	0.08	0.01	<div>Intersection between Japanese legal concepts (J1-J8) and Danish legal concepts (D1-D9)</div> <div>If mean value is equal or over 0.40.43</div> <div>If mean value is equal or over 0.30.34</div> <div>If mean value is equal or over 0.20.21</div> <div>If mean value is equal or over average of all mean values0.25</div> <div>Japanese or Danish feature structures for each legal concept cluster</div> <div>If likelihood is equal of over 0.50.50</div>													
Df2	0.03	0.03	0.09	0.17	0.09	0.95	0.06	0.19	0.11	0.33														
Df3	0.10	0.71	0.07	0.07	0.23	0.07	0.10	0.10	0.17	0.20														
Df4	0.05	0.06	0.07	0.07	0.07	0.08	0.70	0.67	0.17	0.83														
Df5	0.06	0.07	0.09	0.73	0.73	0.09	0.13	0.86	0.20	0.20														
Df6	0.24	0.07	0.36	0.82	0.91	0.91	0.88	0.86	0.20	0.80														
Df7	0.06	0.64	0.55	0.09	0.18	0.91	0.75	0.13	0.20	0.80														
Df8	0.91	0.89	0.88	0.88	0.88	0.88	0.17	0.20	0.25	0.25														
Df9	0.58	0.90	0.88	0.75	0.75	0.13	0.83	0.67	0.25	0.50														
Df10	0.71	0.83	0.60	0.20	0.20	0.20	0.75	0.25	0.33	0.33														

Fig 9
SUBSTITUTE SHEET (RULE 26)

11/33

J1	Upper secondary school, full day specialized course Upper secondary school, day/evening specialized course Upper secondary school, correspondence specialised course Upper secondary school, (full day/evening school), short-term course Secondary education school (upper division), full day specialized course Secondary education school (upper division), day/evening specialized course Secondary education school (upper division), full day, day/evening short-term School for Special Needs Education, upper secondary department, specialized School for Special Needs Education, upper secondary department short-term College of technology, regular course Specialized training college, upper secondary course	Jf4, Jf8, Jf9, Jf11	Jf3	'f16:theoretical cumulative at ISCED 5: medium' 'f17:theoretical cumulative at ISCED 5: long' 'f19:qualification structure at ISCED 5: 1st degree' 'f21:qualification structure at ISCED 5: 2nd degree' 'f23:qualification structure at ISCED 5: BA' 'f24:qualification structure at ISCED 5: MA' 'f32:Minimum entrance requirement (ISCED level) 5A 1st' 'f56:starting age: 22' 'f66:duration: 2' 'f100:cumulative duration: 18' 'f101:cumulative duration: 18-20'
J2	Upper secondary school, full day general course Upper Secondary school, day/evening general course Upper secondary school, correspondence general course Upper secondary school, full day integrated course Upper secondary school, day/evening integrated course Secondary education school (upper division), full day general course Secondary education school (upper division), day/evening general course Secondary education school (upper division), full day integrated course Secondary education school (upper division), day/evening integrated course School for Special Needs Education, upper secondary, general course	Jf4, Jf8, Jf11, Jf13, Jf14	Jf4	'f4:ISCED 3' 'f27:Minimum entrance requirement (ISCED level) 2A' 'f28:Minimum entrance requirement (ISCED level) 2B' 'f42:starting age: 15' 'f86:cumulative duration: 12'
J3	Upper secondary school, (full day, day/evening), advanced course Secondary education (upper division), (full day, day/evening), advanced School for Special Needs Education, upper secondary, advanced course Junior college, short-term course University, short-term course Junior college, regular course Junior college, advanced course College of technology, regular course College of technology, advanced course Specialised training college, post-secondary course	Jf6, Jf8, Jf10	Jf6	'f6:ISCED 5' 'f29:Minimum entrance requirement (ISCED level) 3A' 'f30:Minimum entrance requirement (ISCED level) 3B' 'f47:starting age: 18'
J4	Day nursery Kindergarten School for Special Needs Education, kindergarten department Elementary school School for Special Needs Education, elementary department University, graduate school, Doctor's course University, graduate school, Doctor's course of pharmacy, medicine etc. Specialised training college, general course Miscellaneous schools	Jf8	Jf7	'f3:ISCED 2' 'f26:Minimum entrance requirement (ISCED level) 1' 'f40:starting age: 12' 'f82:cumulative duration: 9'
J5	University, undergraduate University, undergraduate of pharmacy, medicine etc. University, advanced course University, graduate school, Master's course University, professional graduate school University, professional graduate law school	Jf8, Jf13	Jf8	'f109:no work based element' 'f111: not specifically designed for adults ' 'f113:not specifically designed for part-time attendance'
J6	University, undergraduate, correspondence course University, graduate school, Master's course correspondence University, professional graduate school, correspondence	Jf3, Jf8, Jf12, Jf13	Jf9	'f11:programme destination C' 'f14:programme orientation vocational'
J7	Lower secondary school Secondary education school(lower division) School for Special Needs Education, lower secondary department	Jf7, Jf8, Jf11, Jf13, Jf14	Jf10	'f63:duration: 1' 'f64:duration: 1+'
J8	University, graduate school, Doctor's course correspondence	Jf8, Jf12	Jf11	'f70:duration: 3'
J9	Junior college, correspondence course	Jf6, Jf8, Jf12	Jf12	'f112:specifically designed for part-time attendance'
			Jf13	'f9:programme destination A'
			Jf14	'f12:programme orientation General'
D1	Kindergarten Pre-school class in primary school Primary level 1st-6th grade Admittance courses for programmes at 5A and 5B Doctoral programmes	Df8, Df9, Df10	Df2	'f11:programme destination C' 'f14:programme orientation vocational' 'f45:starting age: 16-30' 'f71:duration 3 +' 'f73:duration: 3-5' 'f92:cumulative duration: 14' 'f108:work based element '
D2	Upper secondary higher commercial examination Upper secondary higher technical examination HF higher preparatory examination Upper secondary school leaving examination	Df3, Df7, Df8, Df9, Df10	Df4	'f51:starting age: 18-50' 'f60:duration: 0.2-1'
D3	Lower secondary level 7th-10th grade home economics and needlework Folk- and youth high-school	Df7, Df8, Df9, Df10	Df5	'f110: specifically designed for adults ' 'f112:specifically designed for part-time attendance'
D4	Medium cycle tertiary education Master's programme Undivided Master degrees	Df5, Df6, Df8, Df9	Df6	'f6:ISCED 5' 'f19:qualification structure at ISCED 5: 1st degree' 'f29:Minimum entrance requirement (ISCED level) 3A'
D5	Practical admittance courses for programmes at 5B Short cycle tertiary education Bachelor's programme	Df5, Df6, Df8, Df9	Df7	'f53:starting age: 20' 'f56:starting age: 22' 'f57:starting age: 24' 'f4:ISCED 3'
D6	Upper secondary, health service assistant, vocational education Upper secondary, Carpenter, blacksmith, electrician, vocational education Upper secondary, agriculture, horticulture, forestry, vocational education	Df2, Df6, Df7, Df8	Df8	'f27:Minimum entrance requirement (ISCED level) 2A' 'f47:starting age: 18'
D7	General adult education 9th-10th grade Higher preparatory examination, single subject education	Df4, Df6, Df7, Df9, Df10	Df9	'f111: not specifically designed for adults ' 'f113:not specifically designed for part-time attendance'
D8	Short cycle tertiary education, open education Post secondary tertiary education, open education	Df4, Df5, Df6, Df9	Df10	'f9:programme destination A' 'f109:no work based element '
D9	Doctorate degree			
D10	Upper secondary, open vocational education	Df4, Df6, Df7		

Fig 10

12/33

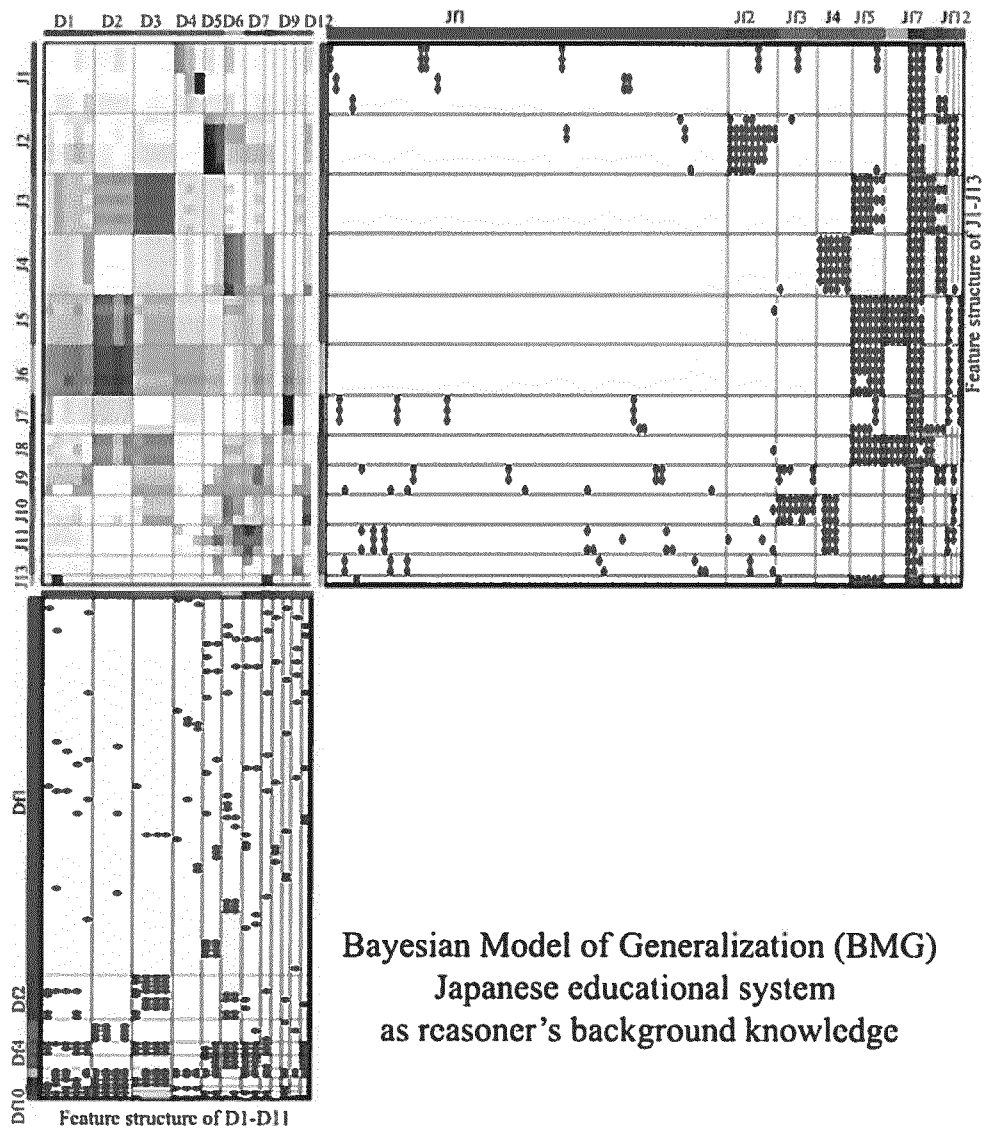


Fig 11

13/33

Avg. 0.15	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12	J1	J2	J3	J4	J5	J6	J7	J8	J9	J10	J11	J12
J1	0.05	0.06	0.03	0.03	0.06	0.05	0.04	0.05	0.00	0.05	0.06	0.08	0.04	0.06	0.08	0.02	0.09	0.03	0.96	0.07	0.47	0.11	0.11	0.11
J2	0.11	0.07	0.09	0.03	0.70	0.20	0.15	0.12	0.00	0.05	0.03	0.19	0.01	0.73	0.08	0.02	0.05	0.04	0.85	0.07	0.21	0.88	0.88	0.13
J3	0.19	0.32	0.59	0.12	0.10	0.06	0.11	0.10	0.00	0.10	0.14	0.10	0.00	0.02	0.02	0.03	0.81	0.04	0.95	0.93	0.50	0.13	0.13	0.13
J4	0.16	0.06	0.12	0.08	0.12	0.51	0.15	0.40	0.00	0.06	0.06	0.19	0.00	0.02	0.05	0.92	0.03	0.04	0.95	0.07	0.92	0.13	0.25	0.13
J5	0.22	0.51	0.22	0.13	0.10	0.19	0.17	0.35	0.00	0.30	0.10	0.05	0.00	0.04	0.03	0.03	0.97	0.95	0.88	0.08	0.08	0.86	0.14	0.86
J6	0.44	0.69	0.27	0.09	0.09	0.09	0.09	0.09	0.00	0.40	0.09	0.13	0.00	0.02	0.03	0.03	0.91	0.05	0.94	0.08	0.08	0.86	0.14	0.86
J7	0.11	0.12	0.05	0.07	0.06	0.02	0.06	0.06	0.00	0.83	0.05	0.03	0.05	0.03	0.03	0.04	0.31	0.06	0.93	0.22	0.22	0.67	0.17	0.67
J8	0.11	0.35	0.40	0.06	0.05	0.18	0.11	0.20	0.00	0.16	0.09	0.16	0.00	0.07	0.04	0.05	0.95	0.93	0.82	0.88	0.13	0.20	0.20	0.20
J9	0.07	0.03	0.19	0.04	0.11	0.36	0.14	0.11	0.00	0.02	0.12	0.32	0.07	0.03	0.30	0.05	0.05	0.07	0.91	0.13	0.63	0.20	0.60	0.20
J10	0.07	0.04	0.05	0.08	0.16	0.32	0.54	0.11	0.00	0.03	0.03	0.24	0.00	0.10	0.83	0.50	0.05	0.07	0.82	0.13	0.13	0.20	0.80	0.20
J11	0.07	0.04	0.05	0.01	0.18	0.06	0.12	0.05	0.36	0.01	0.40	0.05	0.08	0.14	0.04	0.50	0.05	0.07	0.82	0.13	0.13	0.80	0.60	0.20
J12	0.04	0.01	0.05	0.01	0.04	0.06	0.05	0.11	0.00	0.04	0.02	0.05	0.08	0.20	0.06	0.07	0.07	0.10	0.75	0.17	0.17	0.25	0.25	0.25
J13	0.13	0.09	0.05	0.05	0.04	0.06	0.05	0.68	0.00	0.04	0.06	0.13	0.03	0.09	0.22	0.13	0.88	0.17	0.80	0.25	0.50	0.33	0.33	0.50
D1	0.04	0.01	0.01	0.05	0.12	0.10	0.08	0.07	0.08	0.08	0.09	0.08	0.08	0.09	0.08	0.09	0.09	0.03	0.96	0.07	0.47	0.11	0.11	0.11
D2	0.19	0.02	0.67	0.03	0.05	0.41	0.14	0.17	0.08	0.17	0.08	0.17	0.08	0.09	0.08	0.09	0.09	0.03	0.96	0.07	0.47	0.11	0.11	0.11
D3	0.04	0.59	0.05	0.06	0.08	0.08	0.17	0.29	0.14	0.29	0.29	0.14	0.08	0.09	0.08	0.09	0.09	0.03	0.96	0.07	0.47	0.11	0.11	0.11
D4	0.75	0.07	0.93	0.09	0.75	0.88	0.88	0.80	0.20	0.20	0.20	0.80	0.08	0.09	0.08	0.09	0.09	0.03	0.96	0.07	0.47	0.11	0.11	0.11
D5	0.06	0.07	0.07	0.09	0.63	0.88	0.88	0.40	0.20	0.20	0.20	0.80	0.08	0.09	0.08	0.09	0.09	0.03	0.96	0.07	0.47	0.11	0.11	0.11
D6	0.58	0.89	0.70	0.86	0.80	0.17	0.83	0.75	0.25	0.25	0.25	0.75	0.08	0.09	0.08	0.09	0.09	0.03	0.96	0.07	0.47	0.11	0.11	0.11
D7	0.58	0.90	0.90	0.13	0.17	0.17	0.17	0.25	0.25	0.25	0.25	0.33	0.08	0.09	0.08	0.09	0.09	0.03	0.96	0.07	0.47	0.11	0.11	0.11
D8	0.71	0.83	0.83	0.17	0.20	0.25	0.25	0.33	0.33	0.33	0.33	0.33	0.08	0.09	0.08	0.09	0.09	0.03	0.96	0.07	0.47	0.11	0.11	0.11
D9	0.86	0.83	0.83	0.17	0.20	0.25	0.25	0.33	0.33	0.33	0.33	0.33	0.08	0.09	0.08	0.09	0.09	0.03	0.96	0.07	0.47	0.11	0.11	0.11
D10	0.83	0.83	0.83	0.33	0.80	0.75	0.75	0.67	0.33	0.67	0.67	0.67	0.08	0.09	0.08	0.09	0.09	0.03	0.96	0.07	0.47	0.11	0.11	0.11

Intersection between Japanese legal concepts (J1-J8) and Danish legal concepts (D1-D9)

If mean value is equal or over 0.4

If mean value is equal or over 0.3

If mean value is equal or over 0.2

If mean value is equal or over average of all mean values

Japanese or Danish feature structures for each legal concept cluster

If likelihood is equal or over 0.5

Fig 12

Fig 13

15/33

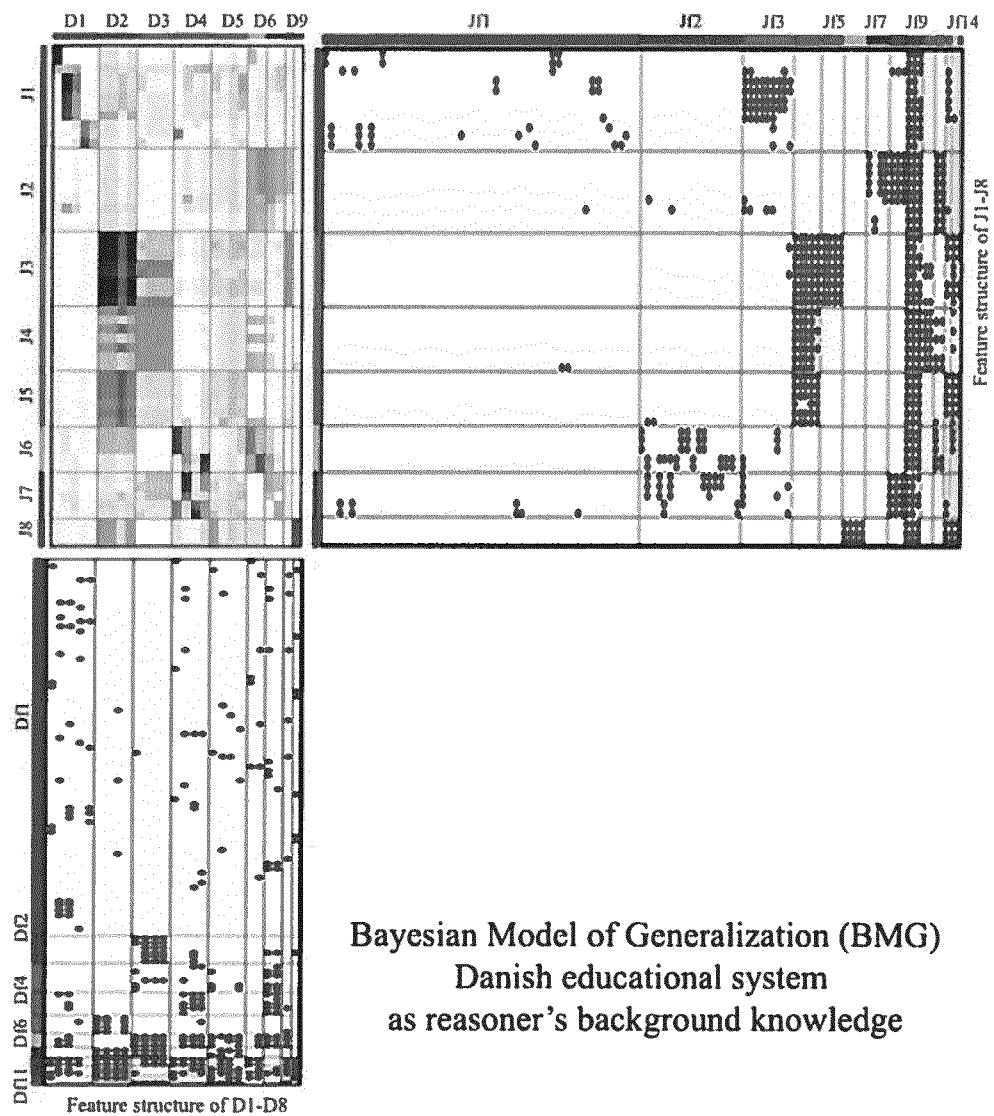


Fig 14

16/33

Avr.0.12	D1	D2	D3	D4	D5	D6	D7	D8	D9	J1	J2	J3	J4	J5	J6	J7	J8	J9	J10	J11	J12	J13	J14
J1	0.19	0.11	0.06	0.08	0.09	0.07	0.06	0.05	0.02	0.06	0.01	0.01	0.46	0.02	0.02	0.02	0.12	0.89	0.04	0.04	0.54	0.15	0.08
J2	0.03	0.08	0.07	0.06	0.06	0.06	0.06	0.13	0.02	0.00	0.02	0.06	0.02	0.03	0.03	0.49	0.66	0.96	0.05	0.95	0.18	0.09	0.10
J3	0.03	0.69	0.06	0.06	0.13	0.07	0.04	0.32	0.05	0.00	0.01	0.04	0.98	0.97	0.03	0.03	0.03	0.04	0.88	0.39	0.06	0.60	0.60
J4	0.02	0.34	0.04	0.04	0.09	0.14	0.05	0.03	0.04	0.01	0.02	0.86	0.03	0.03	0.03	0.03	0.04	0.95	0.94	0.56	0.11	0.38	0.11
J5	0.03	0.43	0.11	0.07	0.14	0.11	0.02	0.06	0.04	0.00	0.03	0.02	0.90	0.04	0.04	0.04	0.05	0.95	0.07	0.08	0.75	0.86	0.71
J6	0.03	0.15	0.06	0.06	0.07	0.34	0.11	0.04	0.02	0.00	0.32	0.13	0.04	0.05	0.05	0.05	0.06	0.94	0.08	0.67	0.14	0.57	0.14
J7	0.03	0.09	0.14	0.12	0.15	0.05	0.16	0.09	0.02	0.03	0.24	0.17	0.04	0.05	0.05	0.05	0.94	0.82	0.08	0.08	0.43	0.14	0.14
J8	0.03	0.06	0.06	0.07	0.15	0.09	0.01	0.05	0.05	0.01	0.02	0.03	0.06	0.07	0.93	0.07	0.09	0.91	0.13	0.13	0.80	0.80	0.80
D1	0.10	0.01	0.01	0.05	0.03	0.06	0.07	0.07	0.10	Intersection between Japanese legal concepts (J1-J8) and Danish legal concepts (D1-D9)													
D2	0.03	0.04	0.81	0.15	0.04	0.07	0.21	0.13	0.25	If mean value is equal or over 0.4													
D3	0.03	0.04	0.27	0.15	0.23	0.07	0.71	0.13	0.13	If mean value is equal or over 0.3													
D4	0.19	0.05	0.05	0.59	0.05	0.08	0.83	0.29	0.14	If mean value is equal or over 0.2													
D5	0.09	0.72	0.06	0.11	0.06	0.10	0.10	0.33	0.17	If mean value is equal or over average of all mean values													
D6	0.38	0.07	0.93	0.69	0.71	0.50	0.88	0.75	0.20	0.12													
D7	0.08	0.90	0.90	0.10	0.70	0.17	0.17	0.25	0.25	Japanese or Danish feature structures for each legal concept cluster													
D8	0.73	0.90	0.70	0.90	0.44	0.80	0.17	0.75	0.67	0.50													
D9	0.58	0.90	0.20	0.56	0.90	0.67	0.67	0.75	0.75	If likelihood is equal or over 0.5													
D10	0.29	0.83	0.17	0.33	0.67	0.75	0.25	0.33	0.67														
D11	0.14	0.17	0.83	0.17	0.83	0.25	0.25	0.67	0.33														

Fig 15

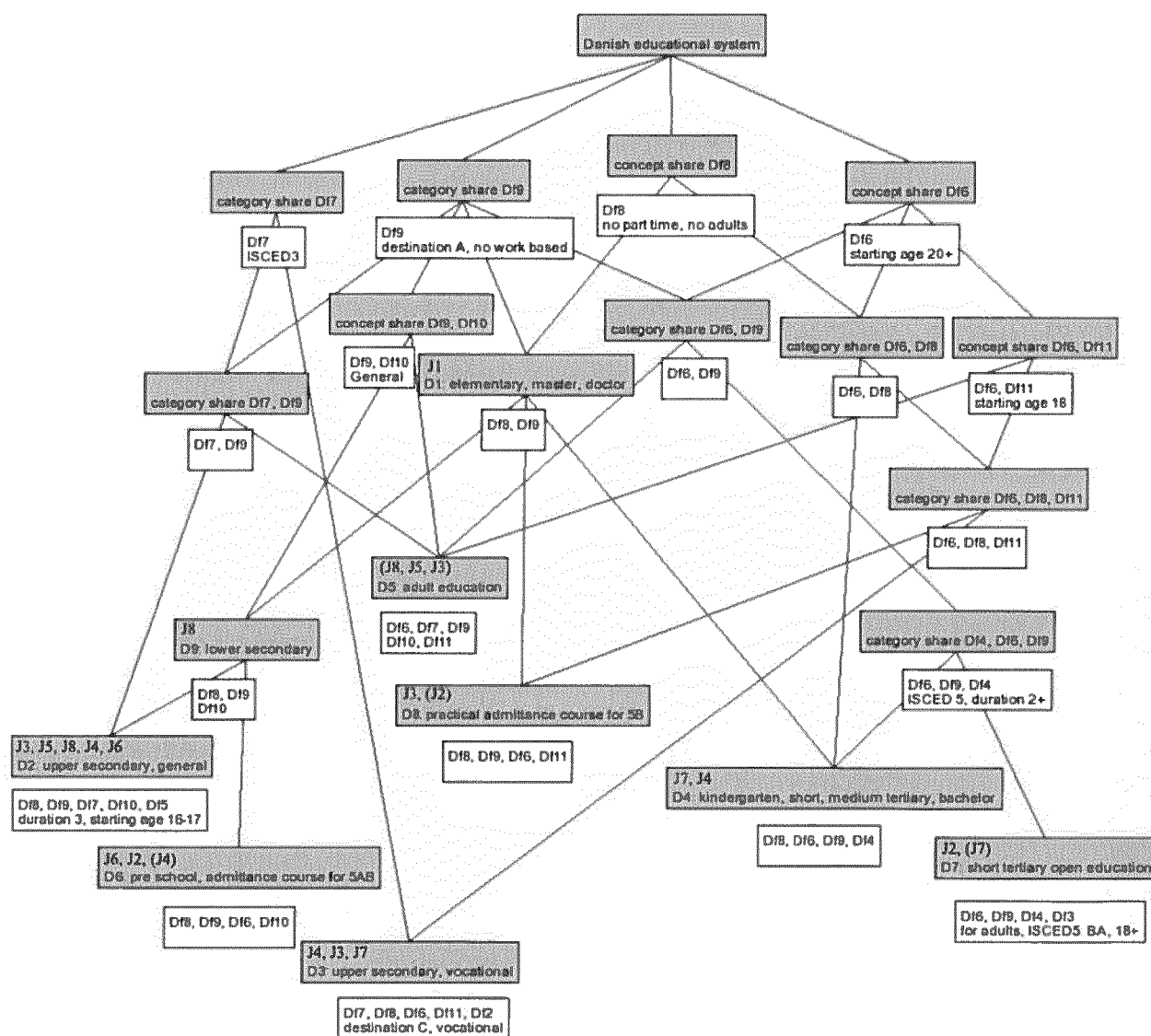
Fig 16

18/33

Primary level 1st-6th grade	0.77	0.77	0.24	0.02	0.02	0.03	0.03	0.03	0.03	0.02
Elementary school										
School for Special Needs Education, elementary department	0.02	0.02	0.20	0.76	0.76	0.62	0.62	0.57	0.03	0.02
University, undergraduate of pharmacy (practical course)	0.02	0.02	0.27	0.46	0.46	0.31	0.31	0.31	0.03	0.11
University, graduate school, Master's correspondence course	0.02	0.02	0.01	0.01	0.02	0.02	0.05	0.42	0.42	0.57
University, professional graduate school, correspondence	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.22	0.13	0.26
University, graduate school, Master's course										
University, professional graduate school										
University, professional graduate law school										
University, graduate school, Doctor's course										
University, graduate school, Doctor's course of pharmacy (practical course)										
University, graduate school, Doctor's correspondence course										
Doctorate	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Doctoral programmes	0.02	0.02	0.01	0.01	0.02	0.02	0.05	0.42	0.42	0.57
Undivided Master degrees	0.02	0.02	0.20	0.76	0.76	0.62	0.62	0.57	0.03	0.02
Master	0.02	0.02	0.20	0.76	0.76	0.62	0.62	0.57	0.03	0.02
Primary level 1st-6th grade	0.77	0.77	0.24	0.02	0.02	0.03	0.03	0.03	0.03	0.02

Fig 17

Fig 19



21/33

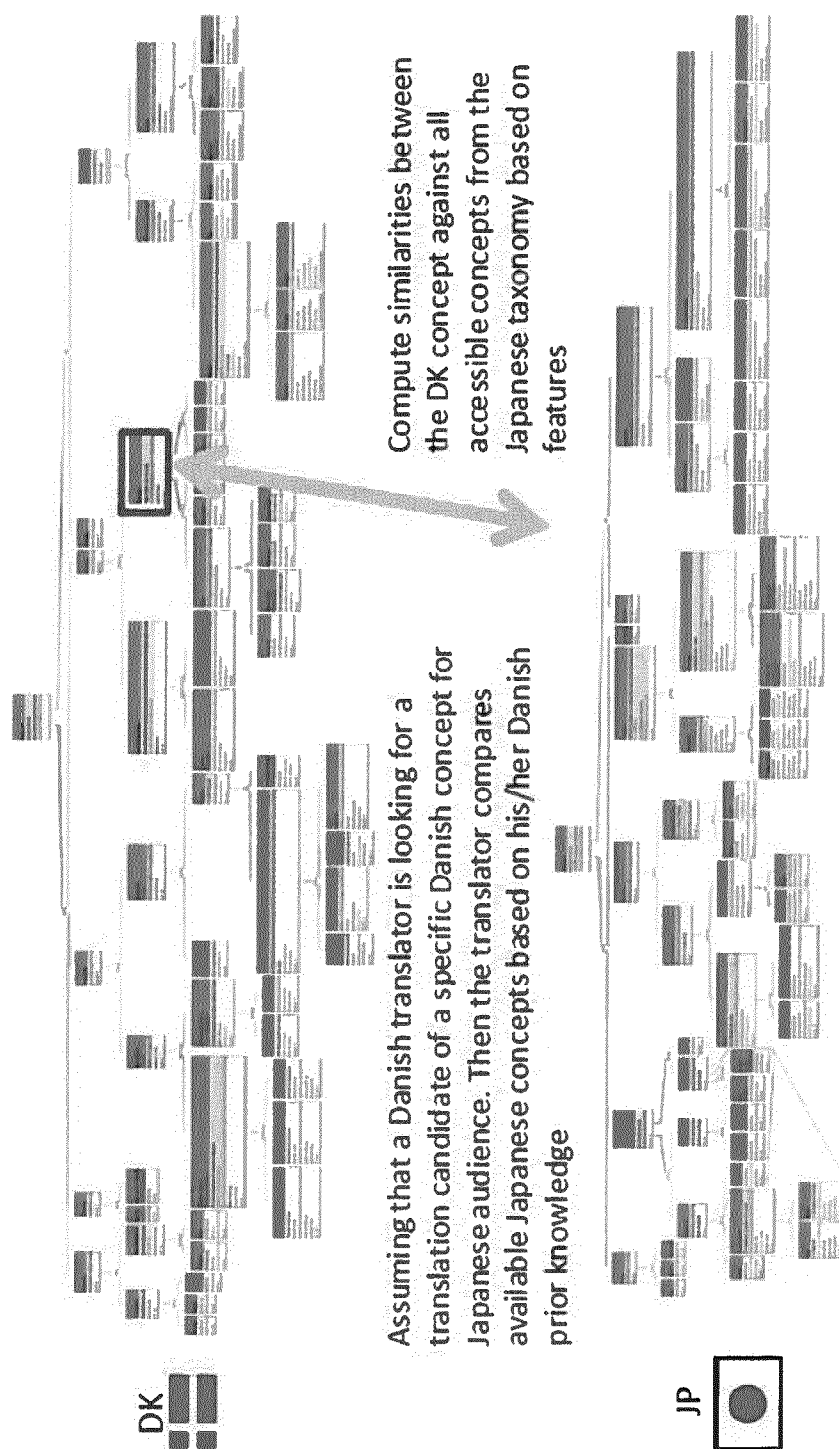


Fig 20a

22/33

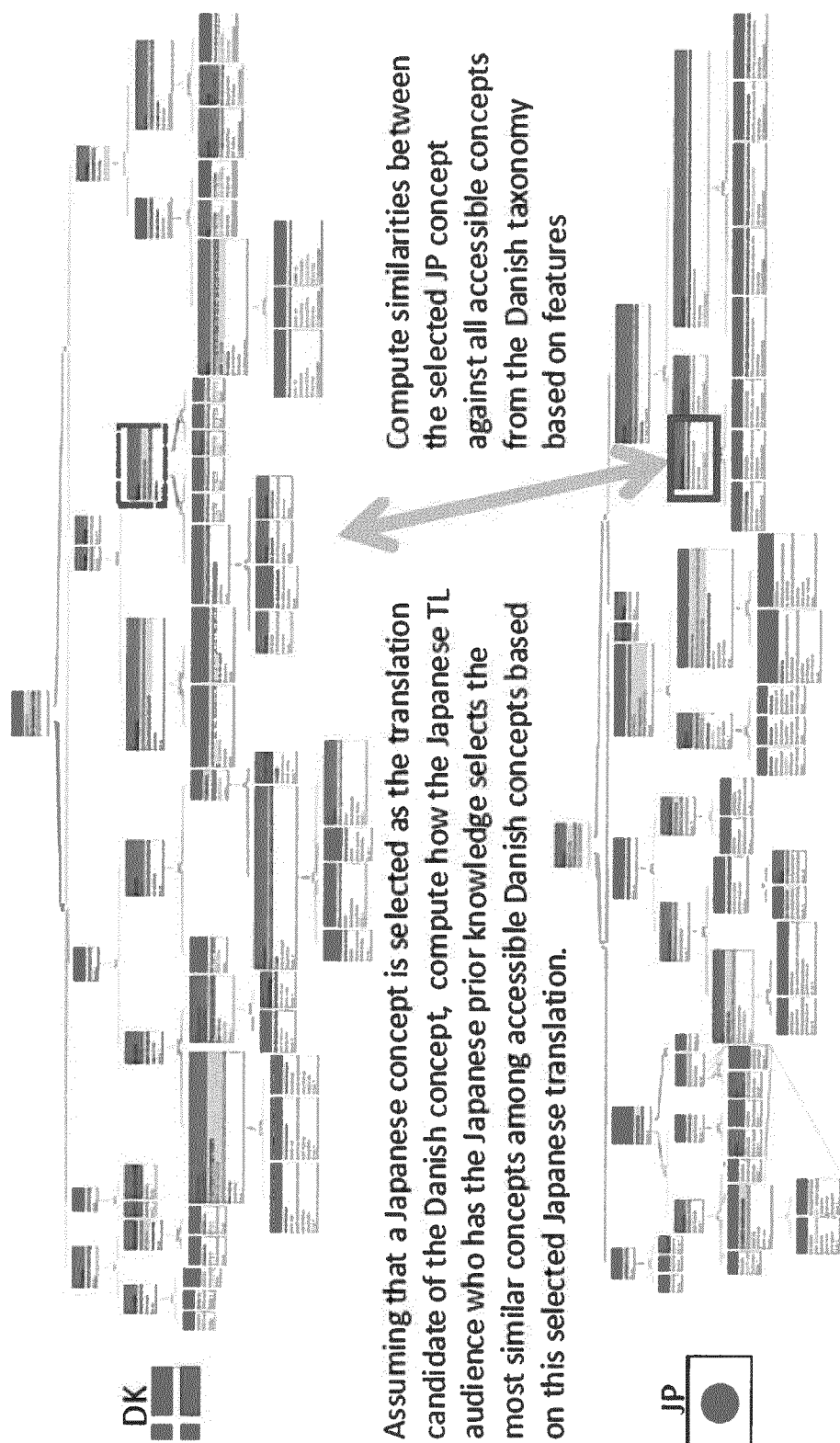


Fig 20b

23/33

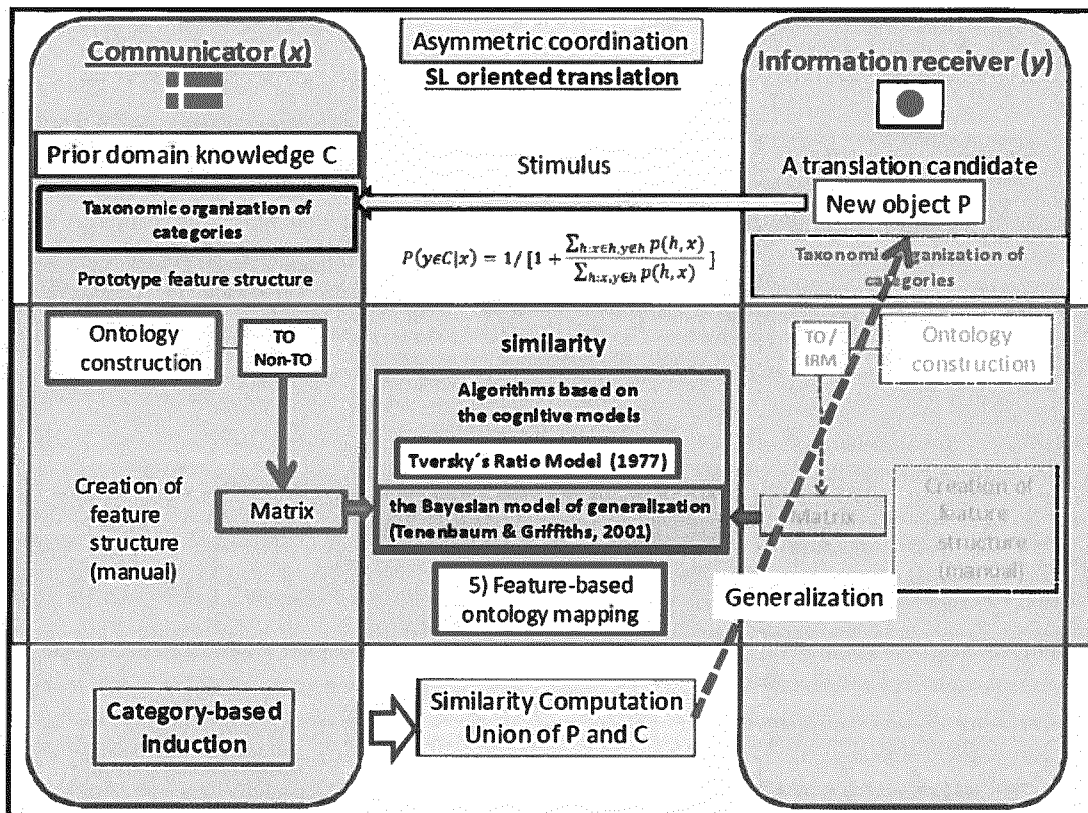


Fig 20c

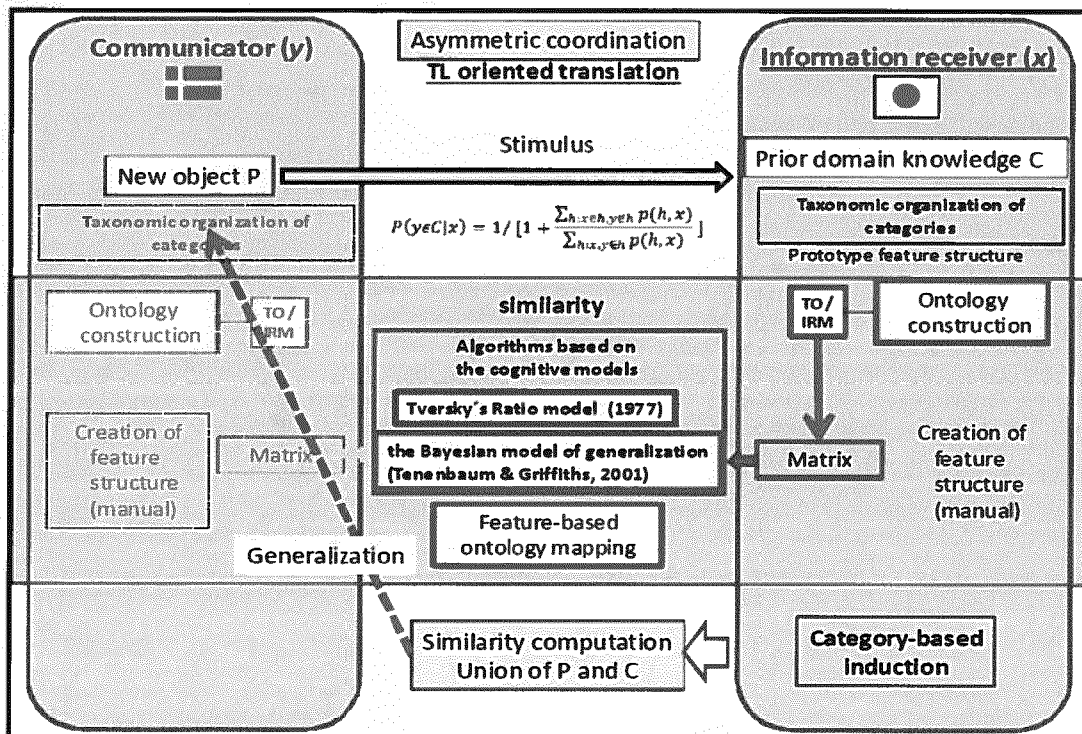


Fig 20d

24/33

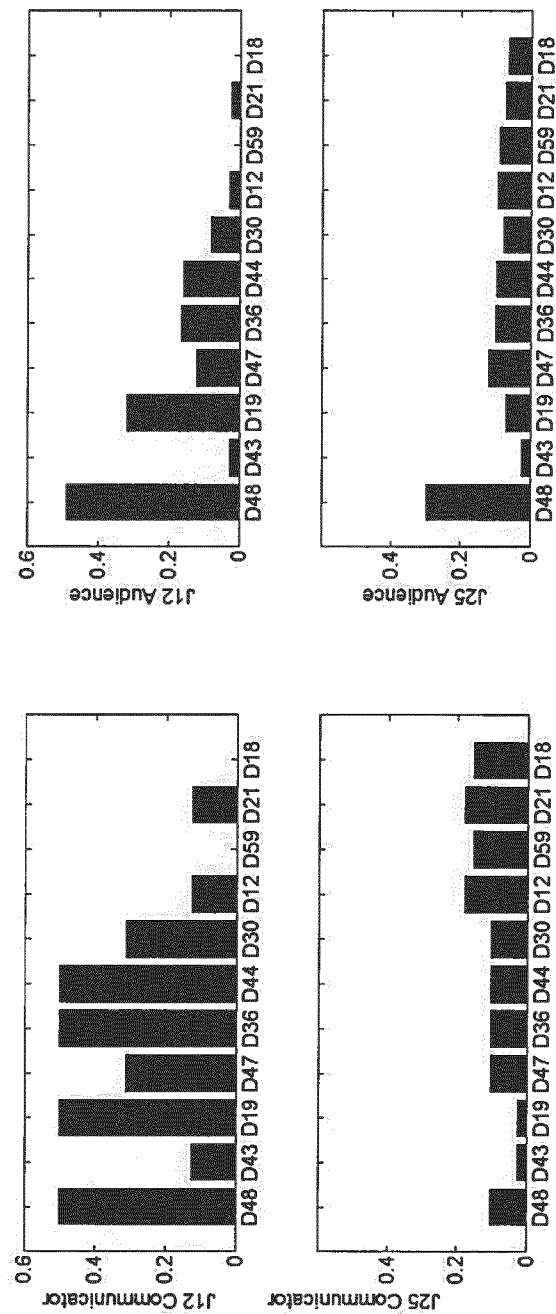


Fig 21a

25/33

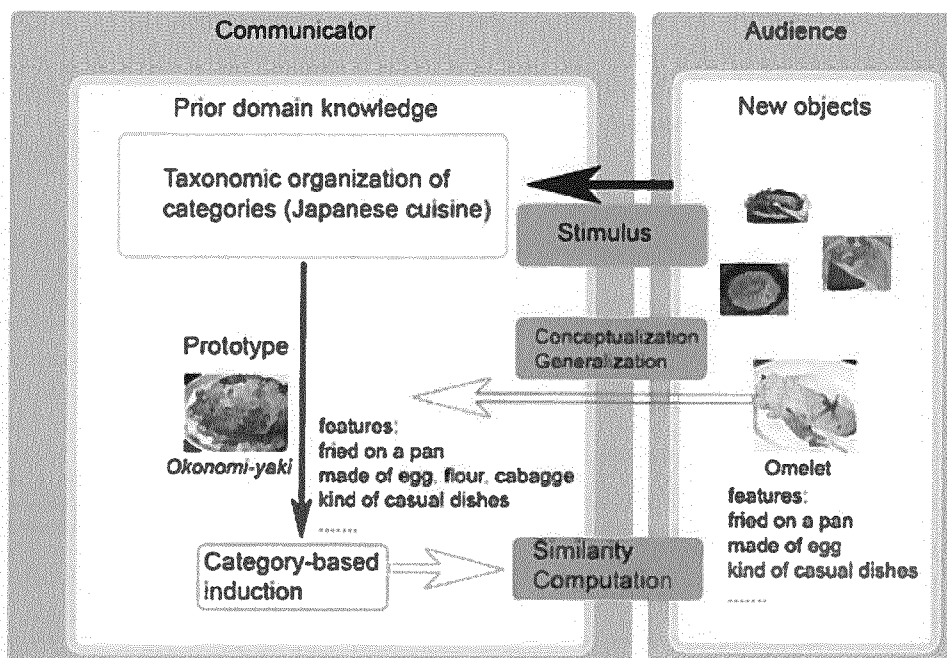


Fig 21b

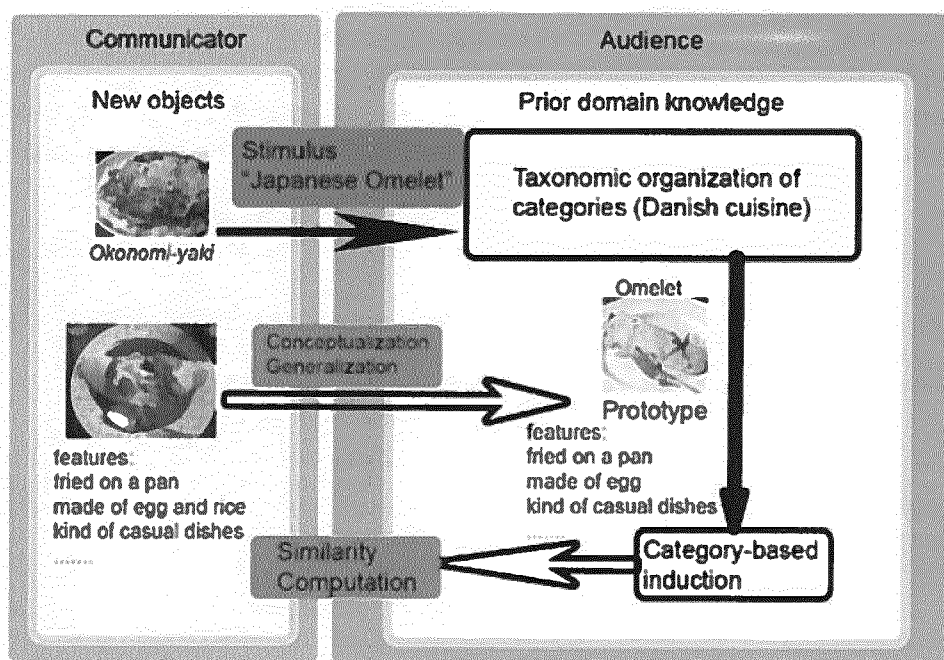


Fig 21c

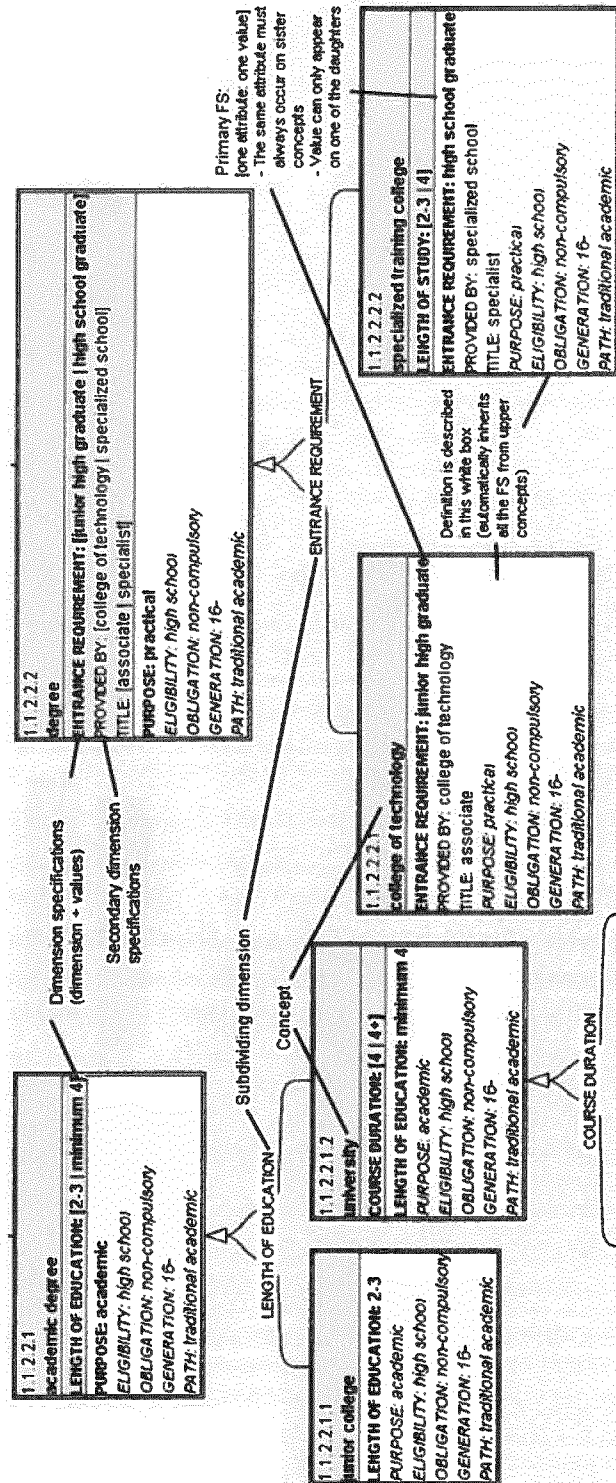


Fig 22

27/33

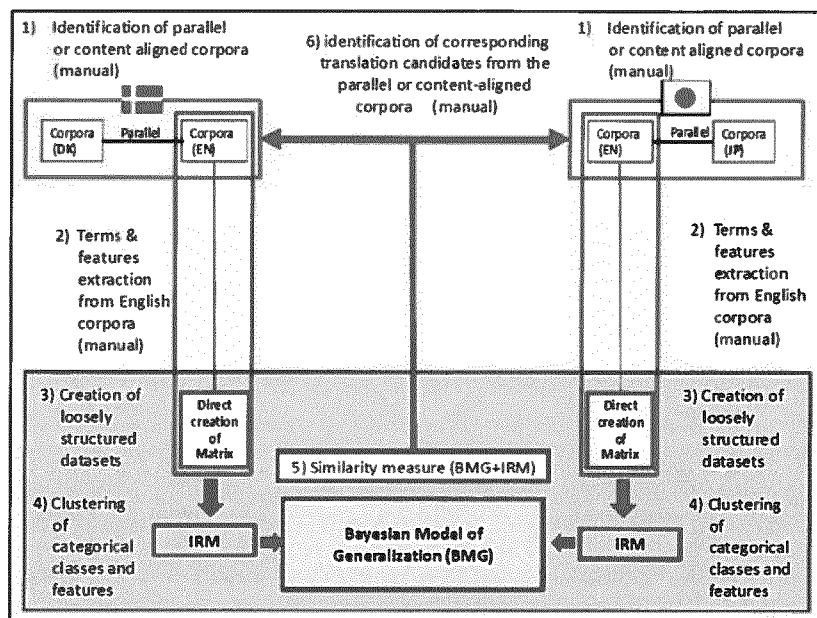


Fig 23a

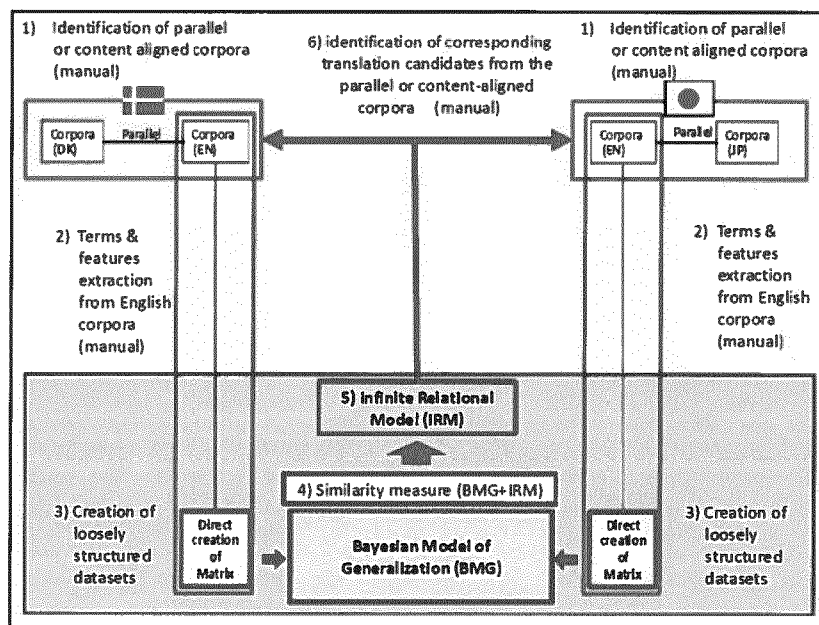


Fig 23b

28/33

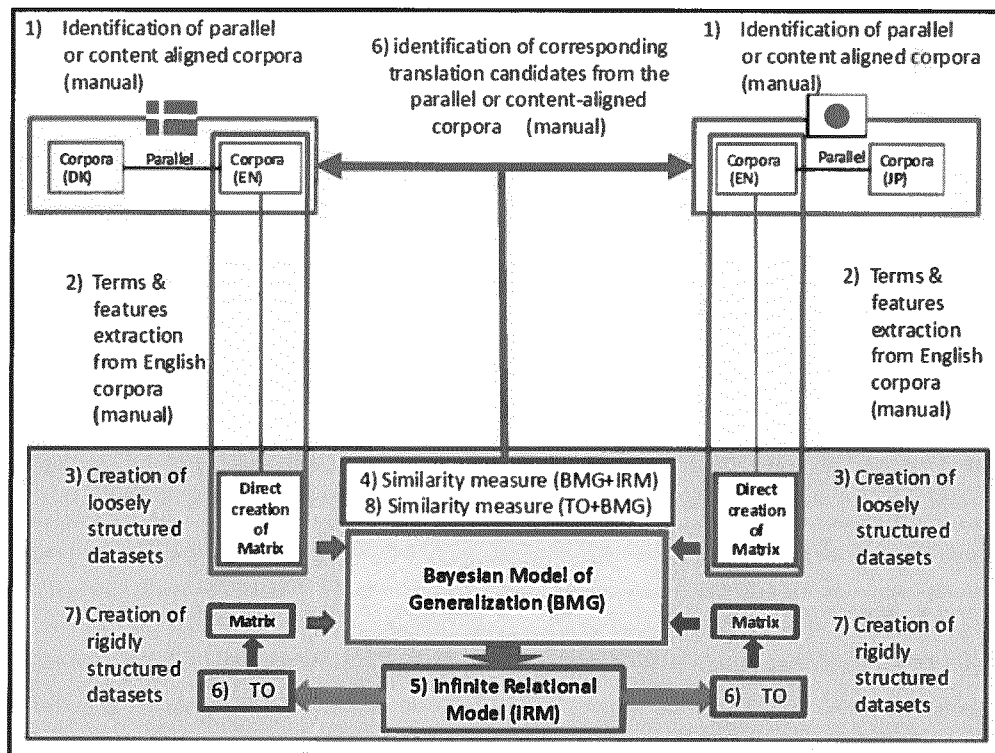


Fig 23c

29/33

Graph sorted according to extracted assignments



Fig 24a

30/33

<p>Cluster 1 (continuing education)</p> <p>D13: continuing education and training for adults</p> <p>D14: continuing vocational education and training</p> <p>D16: Diploma programmes</p> <p>D20: Folk High School</p> <p>D22: further education for adults</p> <p>D23: general adult education</p> <p>D26: higher preparatory single subjects courses</p> <p>D28: higher education for adults</p> <p>D35: master programme</p> <p>D37: non-formal adult education</p> <p>D39: open education</p> <p>D4: Advanced adult program</p> <p>D41: preparatory adult education programme for adults</p> <p>D5: basic adult education</p>	<p>Cluster 3 (higher education)</p> <p>D10: Centres for Higher Education</p> <p>D11: college sector</p> <p>D17: doctoral degree</p> <p>D2: Academies of professional higher education</p> <p>D40: PhD programme</p> <p>D45: professional bachelor programme</p> <p>D50: tertiary education</p> <p>D51: University Bachelor degree programme</p> <p>D52: University college</p> <p>D53: University education</p> <p>D9: Candidates degree programme</p>
<p>Cluster 2 (upper secondary)</p> <p>D24: general upper secondary education</p> <p>D25: gymnasium/STX</p> <p>D27: HHX (højere handels eksamen)</p> <p>D29: higher preparatory exam/HF</p> <p>D31: HTX (højere teknisk eksamen)</p> <p>D34: main course</p> <p>D54: upper secondary education</p> <p>D56: vocational education and training (I-VET)</p> <p>D57: vocational upper secondary education</p> <p>D58: youth education programmes</p> <p>D6: basic course</p> <p>D7: Basic vocational education and training</p>	<p>Cluster 4 (primary and lower secondary)</p> <p>D19: first stage</p> <p>D30: home tuition</p> <p>D36: municipal school/folkeskole</p> <p>D43: preschool class/0th form/børnehave klasse</p> <p>D47: second stage</p> <p>D48: primary and lower secondary education</p>
<p>Cluster 9 (alternative lower secondary)</p> <p>D12: Continuation school (efterskole)</p> <p>D44: private school</p> <p>Cluster 10 (bridge building course)</p> <p>D1: 10th form/bridge building</p> <p>Cluster 11 (lower secondary graduate)</p> <p>D46: school-leaving examinations</p>	<p>Cluster 5 (pre primary)</p> <p>D15: day-care facilities</p> <p>D32: integrated institution</p> <p>D33: Kindergarten (børnehave)</p> <p>D36: Nursery/Creche/vuggestue</p> <p>D42: pre-primary education</p> <p>Cluster 6 (youth school)</p> <p>D18: Dual system</p> <p>D21: Full time system</p> <p>D59: youth school</p>
	<p>Cluster 7 (vocational academy)</p> <p>D3: Academy profession degree programmes</p> <p>D55: Vocational college</p>
	<p>Cluster 8 (vocational college)</p> <p>D49: technical college</p> <p>D8: business college</p>
<p>Cluster 1 (upper secondary)</p> <p>J1: Advanced course (Senkoka)</p> <p>J13: Fisheries course</p> <p>J14: Full time course</p> <p>J16: General course (Honka)</p> <p>J17: Graded course</p> <p>J19: Home-economics course</p> <p>J2: Agriculture course</p> <p>J31: Nursing course</p> <p>J32: Ordinary education course</p> <p>J44: Public upper secondary school</p> <p>J47: Specialized course (Bekka)</p> <p>J48: Specialized education course</p> <p>J50: Technology course</p> <p>J54: Upper secondary school</p> <p>J6: Commerce course</p> <p>J7: Comprehensive education course</p> <p>J8: Correspondence course</p> <p>J9: Credit course</p>	<p>Cluster 2 (higher education)</p> <p>J11: Doctoral degree</p> <p>J18: Graduate school</p> <p>J20: Junior college</p> <p>J26: Masters degree</p> <p>J29: National university</p> <p>J38: Private university</p> <p>J43: Public university</p> <p>J51: Undergraduate department</p> <p>J52: university</p>
<p>Cluster 5 (alternative post compulsory)</p> <p>J15: General course</p> <p>J27: Miscellaneous schools</p> <p>J46: Specialized course</p> <p>J49: Specialized training college</p> <p>J63: Upper secondary course</p> <p>Cluster 6 (special upper secondary + higher education)</p> <p>J3: colleges of technology</p> <p>J34: post-compulsory educational institution</p> <p>J4: colleges of technology</p> <p>– economy, IT management course</p> <p>J5: colleges of technology</p> <p>– industrial course</p>	<p>Cluster 4 (primary)</p> <p>J12: elementary school</p> <p>J25: lower secondary school</p> <p>J28: National elementary school</p> <p>J36: Private elementary school</p> <p>J40: Public elementary school</p> <p>J41: Public lower secondary school</p>
<p>Cluster 7 (secondary)</p> <p>J37: Private six-year secondary school</p> <p>J42: Public six-year secondary school</p> <p>J45: six-year secondary school</p>	<p>Cluster 8 (part time upper secondary)</p> <p>J33: Part time course</p>
	<p>Cluster 9 (professional graduate school)</p> <p>J39: Professional graduate school</p>

Fig 24b

31/33

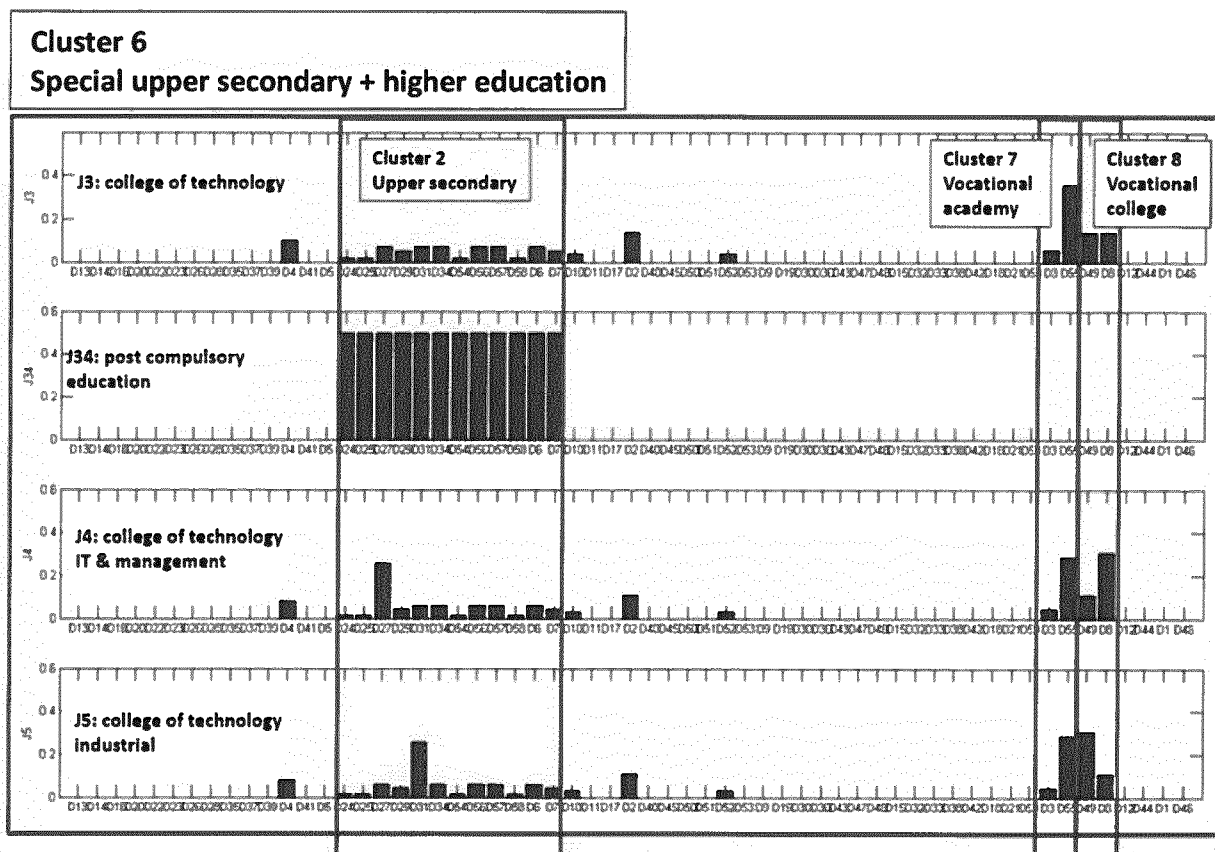


Fig 24c

32/33

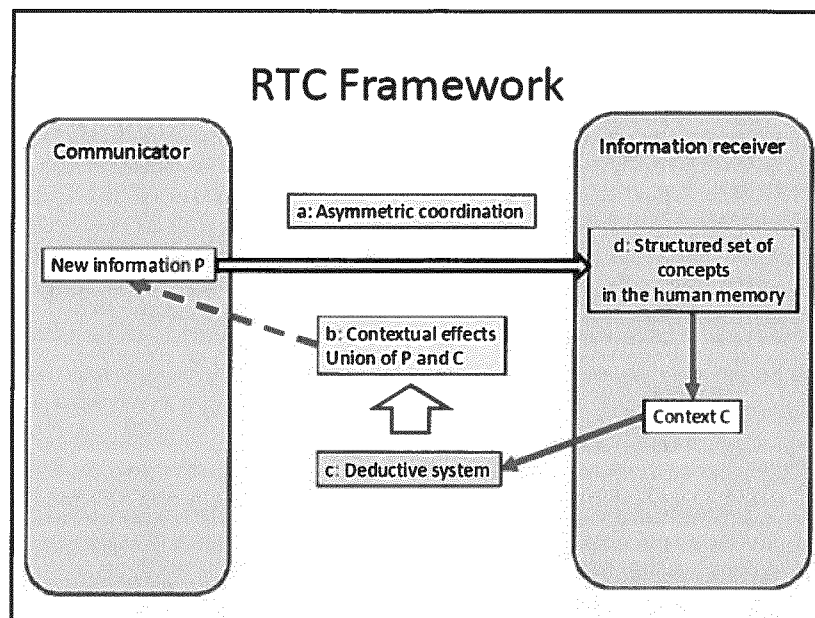


Fig 25a

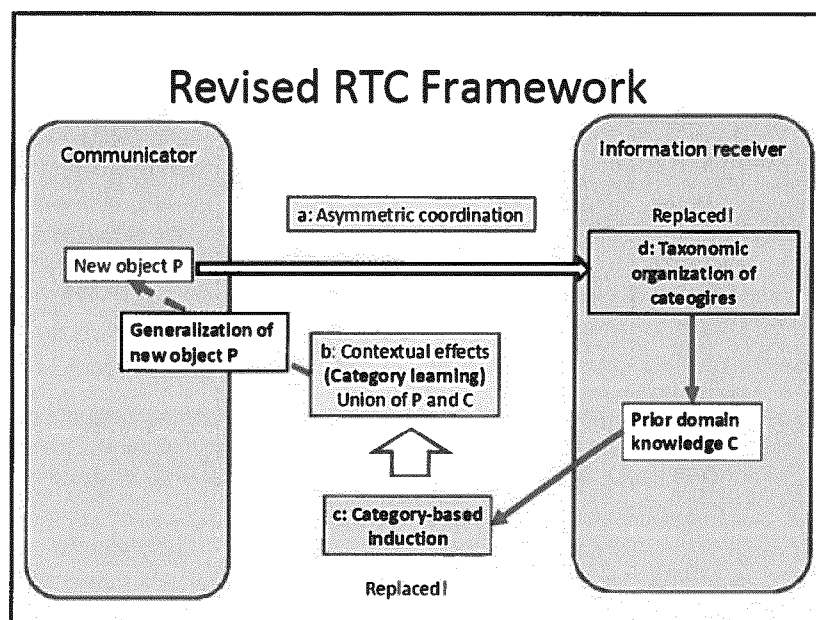


Fig 25b

33/33

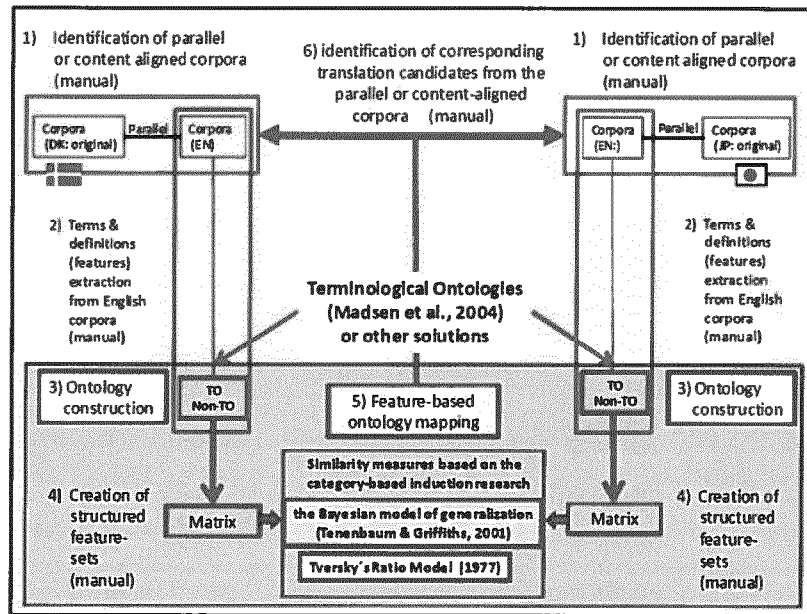


Fig 26

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2013/062040

A. CLASSIFICATION OF SUBJECT MATTER
INV. G06F17/28 G06F17/30
ADD. G06F9/44

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	FUMIKO KANO GLUCKSTAD: "Alignment of Remote Cultures: In Contrast to the Relevance Theory of Communication", CULTURE AND COMPUTING (CULTURE COMPUTING), 2011 SECOND INTERNATIONAL CONFERENCE ON, IEEE, 20 October 2011 (2011-10-20), pages 21-26, XP032074454, DOI: 10.1109/CULTURE-COMPUTING.2011.13 ISBN: 978-1-4577-1593-8	1,2,4, 6-24
Y	the whole document ----- -/--	3,5



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

17 July 2013

Date of mailing of the international search report

24/07/2013

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Woods, Justin

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2013/062040

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	<p>Joshua B. Tenenbaum ET AL: "Generalization, similarity, and Bayesian inference", Behav. Brain Sci. 24, 1 January 2001 (2001-01-01), pages 629-640, XP055045128, Retrieved from the Internet: URL:http://www.cog.brown.edu/courses/cg2000/Papers/Tenenbaum and GriffithsBBS2001.pdf [retrieved on 2012-11-21] cited in the application abstract Section 2</p>	3,5
X	<p>-----</p> <p>Piek Kyoto ET AL: "KYOTO: A System for Mining, Structuring, and Distributing Knowledge Across Languages and Cultures", Proceedings of LREC 2008, 28 May 2008 (2008-05-28), XP055045136, Marrakech, Morocco. Retrieved from the Internet: URL:http://www.vossen.info/docs/2008/KYOTO_LREC2008[6].pdf [retrieved on 2012-11-21] cited in the application abstract Sections 1-4; figure 1</p>	1,15-24
X	<p>-----</p> <p>JASON J JUNG ET AL: "Indirect Alignment between Multilingual Ontologies: A Case Study of Korean and Swedish Ontologies", 3 June 2009 (2009-06-03), AGENT AND MULTI-AGENT SYSTEMS: TECHNOLOGIES AND APPLICATIONS, SPRINGER BERLIN HEIDELBERG, BERLIN, HEIDELBERG, PAGE(S) 233 - 241, XP019117655, ISBN: 978-3-642-01664-6 abstract Section 1; figure 1</p>	20
A	<p>-----</p> <p>TVERSKY A: "FEATURES OF SIMILARITY", PSYCHOLOGICAL REVIEW, AMERICAN PSYCHOLOGICAL ASSOCIATION, WASHINGTON, DC, US, vol. 84, no. 4, 1 July 1977 (1977-07-01), pages 327-352, XP000869970, ISSN: 0033-295X, DOI: 10.1037/0033-295X.84.4.327 cited in the application abstract</p> <p>-----</p> <p style="text-align: center;">-/--</p>	1,3,5,6, 12,15-24

INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2013/062040

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,P	<p>Fumiko Kano Glückstad: "Bridging Remote Cultures: Cross-lingual concept mapping based on the information receiver's prior knowledge", Ph. D Thesis, 24 October 2012 (2012-10-24), XP055045144, Retrieved from the Internet: URL:http://openarchive.cbs.dk/bitstream/handle/10398/8546/Fumiko_Kano_Gluckstad.pdf?sequence=1 [retrieved on 2012-11-21] the whole document</p> <p>-----</p>	1-24