

Title line 1

Title line 2

**A thesis submitted in partial fulfilment
of the requirement for the degree of Doctor of Philosophy**

Name M. Lastname

July 2011

**Cardiff University
School of Computer Science & Informatics**

Declaration

This work has not previously been accepted in substance for any degree and is not concurrently submitted in candidature for any degree.

Signed (candidate)

Date

Statement 1

This thesis is being submitted in partial fulfillment of the requirements for the degree of PhD.

Signed (candidate)

Date

Statement 2

This thesis is the result of my own independent work/investigation, except where otherwise stated. Other sources are acknowledged by explicit references.

Signed (candidate)

Date

Statement 3

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed (candidate)

Date

Copyright © 2011 Name Lastname.

Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.2 or any later version published by the Free Software Foundation; with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts. A copy of the license is included in the section entitled “GNU Free Documentation License”.

A copy of this document in various transparent and opaque machine-readable formats and related software is available at <http://yourwebsite>.

**To People you care
for their patience and support.**

Abstract

We produce interpretable representations, and demonstrate their applicability in interpretable classifiers. Our approach is model-agnostic, given a similarity-based representation, we are able to produce a representation in terms of domain knowledge. We evaluate the interpretability of our representation and provide examples of interpretable classifiers with our representation.

Acknowledgements

Contents

Abstract	iv
Acknowledgements	v
Contents	vi
List of Publications	viii
List of Figures	ix
List of Tables	x
List of Algorithms	xi
List of Acronyms	xii
1 Introduction	1
1.1 General Introduction	1
1.2 Thesis Overview	1
2 Background	2

3	Converting Vector Spaces into Interpretable Representations	3
3.0.1	Related Work	5
3.0.2	Method	5
3.0.3	Quantitative Results	9
3.0.4	Interpretability Results	12
4	Fine-tuning Vector Spaces to Improve Their Directions	13
5	Investigating Neural Networks In Terms Of Directions	14
	GNU Free Documentation License	15
	Bibliography	24

List of Publications

The work introduced in this thesis is based on the following publications.

-
-

List of Figures

List of Tables

List of Algorithms

List of Acronyms

GPS Global Positioning System

Chapter 1

Introduction

1.1 Interpretability

1.2 Representations

1.3 Thesis Overview

Chapter 2

Background

Chapter 3

Converting Vector Spaces into Interpretable Representations

On the web there is a large volume of raw text data, e.g. Reviews of products, movies, anime, books, music, social posts by individuals, self-descriptive text about a website or product, and so on. These are categorized into domains; each domain has its own quirks, knowledge, and method of being brought about. Although a movie review may sound similar to a book review, they typically differ hugely in the distribution of words used.

One approach to making sense of these domains is to produce rules from expert knowledge. An expert in movies would tell you that if the review talks about it being a "cannibal horror film", we can understand that it is likely a scary movie and is related to the original 'Cannibal Holocaust' movie. Encoding this kind of knowledge is difficult, time-consuming, and hard to automate reliably.

Most successful approaches in recent times, like vector-spaces, word-vectors, and others, rely on the distributional model of semantics. This model relies on encoding unstructured text e.g. of a movie review, as a vector, where each dimension corresponds to how frequent each word is, we are able to calculate how similar the entities are, e.g. we know that if two movies have a similar distribution of words in their reviews, like frequent use of the word 'scary', or 'horror', then they would have a higher similarity value. These models, also known as 'semantic spaces' encode this similarity

information spatially.

A domain-specific semantic space contains a lot of domain knowledge that is not made explicit. One approach towards understanding how these spaces represent information is a conceptual space [?]. In this space, we can understand domain entities, e.g. movies in a domain of IMDB movie reviews, to be represented as points, and domain properties to be in regions around these points. The relationships between these regions have been formalized in terms of similarity reasoning and commonsense reasoning.

In this chapter, we focus on further experimenting with one relationship that was formalized in [?]: a ranking of entities on properties. In particular, we use this method of building a representation of entities as a way to convert a vector space into an interpretable representation, for use in an interpretable classifier. The reason that we chose this representation to expand on is because by representing each entity e with a vector v that corresponds to a ranking r , the meaning of each dimension is distinct, and we are able to find labels composed of clusters of words for these dimensions. Here, we make the distinction between a property and a word, a property is a natural property of the space that exists in terms of a ranking of entities, and words are the labels we use to describe this property.

Having distinct, labelled dimensions are beneficial when producing an interpretable classifier, e.g. a Decision Tree classifier that uses this representation: As each ranking being distinct, each node of the tree acts as a step in the reasoning of why an entity was classified a certain way, and that node is appropriately labelled so that we are able to understand it. We show an example tree in ?? Interestingly, we find that if a property is relevant enough to a classification task, classifying using a single ranking performs better than using multiple rankings - as these rankings generalize to new entities.

The main focus of our quantitative experimentation is to demonstrate that the rankings our representation is composed of correspond to domain knowledge. To do so, we use Decision Trees as described in section ??, which we constrain by depth, forcing the classifier to use only a limited amount of dimensions. We can understand our approach

to be suitable if the classifier performs well with use of only a few properties, as we know that these properties are able to represent important concepts in the domain. We compare these to the original representations, which are typically less disentangled, as well as to topic models, another approach towards representing the domain.

We also evaluate the method qualitatively and in terms of interpretability, specifically evaluating the ability of classifiers to appropriately describe how they concluded that a particular entity belongs to a class. The qualitative investigation focuses on demonstrating how different parameters of the method affect the representation through use of examples and analysis.

This chapter continues as follows: We begin by describing the work related to this method, giving valuable context for the utility and potential of our approach. This is followed by an explanation of the method, including the variations we have adopted for our experimental work. We follow this with our qualitative experimentation, explaining how these variations affect the results, as well as the interpretability of the method, and we end with a quantitative analysis on how well we can represent domain knowledge using decision trees constrained to a limited depth.

3.0.1 Related Work

There are many approaches towards improving document representations.

Probabilistic approaches like topic models.

Matrix factorization approaches.

3.0.2 Method

The goal of this method is to obtain a representation composed of salient properties, starting with a domain-specific vector space S_e and its associated bag-of-words (BOW)

representation B_w . To obtain these properties, we use a variant of the unsupervised method proposed in [?], which we explain in this section.

Rankings entities on words

We can understand that only some words will be properties, as only some correspond to domain knowledge, e.g. in a domain of IMDB movies, the word "the" does not correspond to a property of the domain, but the word "horror" does. Initially, we obtain rankings of entities for each word in the space.

As an initial filtering step, we remove words that do not meet a frequency threshold, with the understanding that words that do not occur in a minimum amount of documents are unlikely to correspond to properties as they are too specific to a subset of movies, which would make them difficult to learn. This leaves us with w_n words. We show the kind of words that would be poorly represented in ??.

Then, for each considered word w , a logistic regression classifier is trained to find a hyperplane H_w in the space that separates entities e which contain w in their BOW B_e representation from those that do not. The vector v_w perpendicular to this hyperplane is then taken as a direction that models the word w . In ??, we show an example of this in a toy domain. To rank the objects on the entity, if e is the representation of an entity in the given vector space S_e then we can think of the dot product $v_w \cdot e$ as the value $r_e w$ of object e for vector v_w , and in particular, we take $r_{e1} < r_{e2}$ to mean that e_2 has the property labelled with the word w to a greater extent than e_1 . The result of this is shown in ??. Example entities, with their associated highest and lowest ranking properties, are shown in ??.

Filtering directions to obtain salient properties

With the rankings R_r , we could create a representation of each entity Se , composed of w_n dimensions, where each dimension is a ranking of the entity e on that word $w_r e$.

However, many of the words do not correspond to salient properties. In-order to filter these words out, we evaluate them using a scoring metric, and remove the words that are not sufficiently well scored. We use three different metrics:

Classification accuracy. Evaluating the quality in terms of the accuracy of the logistic regression classifier: if this classifier is sufficiently accurate, it must mean that whether word w relates to object o (i.e. whether it is used in the description of o) is important enough to affect the semantic space representation of o . In such a case, it seems reasonable to assume that w describes a salient property for the given domain.

Cohen’s Kappa. One problem with accuracy as a scoring function is that these classification problems are often very imbalanced. In particular, for very rare words, a high accuracy might not necessarily imply that the corresponding direction is accurate. For this reason, X proposed to use Cohen’s Kappa score instead. In our experiments, however, we found that accuracy sometimes yields better results, so we keep this as an alternative metric.

Normalized Discounted Cumulative Gain This is a standard metric in information retrieval which evaluates the quality of a ranking w.r.t. some given relevance scores [?]. In our case, the rankings r_e of the entity e are those induced by the dot products $v_w \cdot e$ and the relevance scores are determined by the Pointwise Positive Mutual Information (PPMI) score $ppmi(w, e)$, of the word w in the BoW representation of entity e where $ppmi(w, e) = \max\left(0, \log\left(\frac{p_{we}}{p_{w*} \cdot p_{*o}}\right)\right)$, and

$$p_{wo} = \frac{n(w, o)}{\sum_{w'} \sum_{o'} n(w', o')}$$

where $n(w, e)$ is the number of occurrences of w in the BoW representation of object e , $p_{w*} = \sum_{e'} p_{we'}$ and $p_{*e} = \sum_{w'} p_{w'e}$.

In principle, we may expect that accuracy and Kappa are best suited for binary features, as they rely on a hard separation in the space between objects that have the word in their BoW representation and those that do not, while NDCG should be better suited for gradual features. In practice, however, we could not find such a clear pattern in the

differences between the words chosen by these metrics despite often finding different words. In Table ??, we show examples of properties scored highly for each domain.

Clustering salient properties

Although we are able to find the words that are most salient, the properties in the domain may not correspond directly to words. Further, the properties may not be well described by their associated word. In-order to find better representations of properties, we cluster together similar vectors v_w , following the assumption that those vectors which are similar are representing some property more general than their individual words, and we can find it between them. As the final step, we cluster the best-scoring candidate feature directions v_w . Each of these clusters will then define one of the feature directions to be used in applications. The purpose of this clustering step is three-fold: it will ensure that the feature directions are sufficiently different (e.g. in a space of movies there is little point in having *funny* and *hilarious* as separate features), it will make the features easier to interpret (as a cluster of terms is more descriptive than an individual term), and it will alleviate sparsity issues when we want to relate features with the BoW representation, which will play an important role for the fine-tuning method described in the next section.

As input to the clustering algorithm, we consider the N best-scoring candidate feature directions v_w , where N is a hyperparameter. To cluster these N vectors, we have followed the approach proposed in [?], which we found to perform slightly better than K -means. The main idea underlying their approach is to select the cluster centers such that (i) they are among the top-scoring candidate feature directions, and (ii) are as close to being orthogonal to each other as possible. We refer to [?] for more details. The output of this step is a set of clusters C_1, \dots, C_K , where we will identify each cluster C_j with a set of words. We will furthermore write v_{C_j} to denote the centroid of the directions corresponding to the words in the cluster C_j , which can be computed as $v_{C_j} = \frac{1}{|C_j|} \sum_{w_l \in C_j} v_l$ provided that the vectors v_w are all normalized. These centroids

v_{C_1}, \dots, v_{C_k} are the feature directions that are identified by our method.

Table ?? displays some examples of clusters that have been obtained for three of the datasets that will be used in the experiments, modelling respectively movies, place-types and newsgroup postings. For each dataset, we used the scoring function that led to the best performance on development data(see Section ??). Only the first four words whose direction is closest to the centroid v_C are shown. **K-Means Derrac’s K-Means Variation Mean-shift Hdbscan**

3.0.3 Quantitative Results

We demonstrate the effectiveness of our approach on five datasets, each with their associated tasks.

- The IMDB Movie Dataset: 15,000 movies represented by aggregated reviews. On the tasks of Movie Genres, 100 IMDB Keywords, and UK + US Age Certificates.
- Flickr Place-types: 1,383 place-types. On the tasks of three different place-types, Foursquare, Geonames and OpenCYC.
- The 20-Newsgroups dataset: 18,846 newsgroup posting in 20 different categories. On the task of identifying which of the 20 categories the posting is from.
- The IMDB Sentiment Dataset: 50,000 movie reviews, with binary tags for either positive or negative. On the task of identifying if the review is positive or negative.
- The Reuters Dataset: 10655 News articles. On the task of identifying the category of the article.

To test the ability of the identified directions to accurately represent domain concepts in a ranking, we use low-depth decision-trees. Although these classifiers are not inten-

ded to be competitive with more complex classifiers like unbounded decision trees or SVM's, we find that our rankings are sometimes able to outperform these approaches using only a single decision node (equivalent to finding the best cutoff in a single ranking for classification). We use the F1 metric for our experiments, as almost all classes in each dataset are unbalanced.

We obtain the unsupervised representations as follows:

- To obtain the PPMI representation, we first remove terms from the corpus that do not appear at least $(\text{length of the corpus} * 0.01)$ documents, and then trim all entities that were exclusively composed of these terms. We then we apply standard PPMI based on the implementation from svdmi ¹.
- We obtain the MDS spaces for the movies, place-types and wines datasets from the data made available by [?], to obtain the MDS spaces for the other datasets, we use the same method as [?] and using default parameters for the MDSJ library. For each domain, we obtain 20, 50, 100 and 200 dimensional spaces.
- So that we can use the sparse PPMI matrices when obtaining the space, we the TruncatedSVD method from scikit-learn method² with default parameters. For each domain, we obtain 20, 50, 100 and 200 dimensional spaces.
- For the averaged word-vectors (AWV) and the weighted averaged word vectors (AWVw), we average the glove 6B word-vectors³ obtained from the Wikipedia 2014 + Gigaword 5 corpuses. As these are only available in size 50, 100 and 200, and there are not many other commonly used pre-trained word-vectors that offer multiple dimension sizes, differing from other methods we only obtain AWV and AWVw representations of size 50, 100 and 200. As these dimension sizes are hyper-parameters, we can consider average word vectors to be disadvantaged

¹<https://github.com/Bollegala/svdmi/blob/master/src/svdmi.py>

²<https://scikit-learn.org/stable/modules/generated/sklearn.decomposition.TruncatedSVD.htm>

³<https://nlp.stanford.edu/projects/glove/>

on some tasks, but as it is unlikely that there is too much benefit in training our own word-vectors from the relatively small domains, we opted to simplify the process and simply remove this as a hyper-parameter for this method, as well as the averaged method. The averaged word vectors are obtained by multiplying the vectors by the PPMI values, and finding the weighted average of all vectors multiplied in this way.

- For the Doc2Vec vectors, we use hyperparameter optimization to select the appropriate parameters, as the quality of the end space is typically reliant on well-tuned hyperparameters for the dataset. We use [?] as a guideline for which parameters to optimize, re-using the parameters that stayed constant for both their datasets in their tests, specifically the dbow method, glove6B pre-trained 300-dimensional word-vectors, training those word vectors while training the representation, a sub-sampling of 10(-5), and a negative sample of 5. We tune and select between the values of the window size 5, 15, 50, the minimum frequency count 1, 5, 20, and epoch size 20, 40, 100, and as in the other methods, we obtain vectors of size 20, 50, 100, 200, but the hyperparameters for each of these is found individually, as the different space sizes are later evaluated on how well they can produce good directions. We evaluate the quality of the space using a Gaussian SVM on a selected task for each dataset, in the case of Reuters, Newsgroups and Sentiment, we use their associated tasks, for Movies we use the Genres task and Place-types the Foursquare task, as these tasks represent essential concepts in the domain.

Table 1 shows how well unsupervised representations perform. Topic models are included to demonstrate the difference between other simple and interpretable approaches, and Random Forest's are included to demonstrate the difference between our simple but interpretable approach and a model that typically performs well at the task [?], but is difficult to interpret.

Table 2 demonstrates the difference between unsupervised representations and sali-

ent properties, and Table 3 demonstrates the difference between salient properties and clustered salient properties.

3.0.4 Interpretability Results

Chapter 4

Fine-tuning Vector Spaces to Improve Their Directions

Chapter 5

Investigating Neural Networks In Terms Of Directions

GNU Free Documentation License

Version 1.2, November 2002

Copyright © 2000, 2001, 2002 Free Software Foundation, Inc.
59 Temple Place, Suite 330, Boston, MA 02111-1307 USA

Everyone is permitted to copy and distribute verbatim copies of this license document, but changing it is not allowed.

0. Preamble

The purpose of this License is to make a manual, textbook, or other functional and useful document *free* in the sense of freedom: to assure everyone the effective freedom to copy and redistribute it, with or without modifying it, either commercially or non-commercially. Secondly, this License preserves for the author and publisher a way to get credit for their work, while not being considered responsible for modifications made by others.

This License is a kind of “copyleft”, which means that derivative works of the document must themselves be free in the same sense. It complements the GNU General Public License, which is a copyleft license designed for free software.

We have designed this License in order to use it for manuals for free software, because free software needs free documentation: a free program should come with manuals providing the same freedoms that the software does. But this License is not limited to software manuals; it can be used for any textual work, regardless of subject matter or whether it is published as a printed book. We recommend this License principally for works whose purpose is instruction or reference.

1. Applicability and Definitions

This License applies to any manual or other work, in any medium, that contains a notice placed by the copyright holder saying it can be distributed under the terms of

this License. Such a notice grants a world-wide, royalty-free license, unlimited in duration, to use that work under the conditions stated herein. The “Document”, below, refers to any such manual or work. Any member of the public is a licensee, and is addressed as “you”. You accept the license if you copy, modify or distribute the work in a way requiring permission under copyright law.

A “Modified Version” of the Document means any work containing the Document or a portion of it, either copied verbatim, or with modifications and/or translated into another language.

A “Secondary Section” is a named appendix or a front-matter section of the Document that deals exclusively with the relationship of the publishers or authors of the Document to the Document’s overall subject (or to related matters) and contains nothing that could fall directly within that overall subject. (Thus, if the Document is in part a textbook of mathematics, a Secondary Section may not explain any mathematics.) The relationship could be a matter of historical connection with the subject or with related matters, or of legal, commercial, philosophical, ethical or political position regarding them.

The “Invariant Sections” are certain Secondary Sections whose titles are designated, as being those of Invariant Sections, in the notice that says that the Document is released under this License. If a section does not fit the above definition of Secondary then it is not allowed to be designated as Invariant. The Document may contain zero Invariant Sections. If the Document does not identify any Invariant Sections then there are none.

The “Cover Texts” are certain short passages of text that are listed, as Front-Cover Texts or Back-Cover Texts, in the notice that says that the Document is released under this License. A Front-Cover Text may be at most 5 words, and a Back-Cover Text may be at most 25 words.

A “Transparent” copy of the Document means a machine-readable copy, represented in a format whose specification is available to the general public, that is suitable for revising the document straightforwardly with generic text editors or (for images composed of pixels) generic paint programs or (for drawings) some widely available drawing editor, and that is suitable for input to text formatters or for automatic translation to a variety of formats suitable for input to text formatters. A copy made in an otherwise Transparent file format whose markup, or absence of markup, has been arranged to thwart or discourage subsequent modification by readers is not Transparent. An image format is not Transparent if used for any substantial amount of text. A copy that is not “Transparent” is called “Opaque”.

Examples of suitable formats for Transparent copies include plain ASCII without markup, Texinfo input format, \LaTeX input format, SGML or XML using a publicly available DTD, and standard-conforming simple HTML, PostScript or PDF designed for human modification. Examples of transparent image formats include PNG, XCF and JPG. Opaque formats include proprietary formats that can be read and edited only by proprietary word processors, SGML or XML for which the DTD and/or processing tools

are not generally available, and the machine-generated HTML, PostScript or PDF produced by some word processors for output purposes only.

The “Title Page” means, for a printed book, the title page itself, plus such following pages as are needed to hold, legibly, the material this License requires to appear in the title page. For works in formats which do not have any title page as such, “Title Page” means the text near the most prominent appearance of the work’s title, preceding the beginning of the body of the text.

A section “Entitled XYZ” means a named subunit of the Document whose title either is precisely XYZ or contains XYZ in parentheses following text that translates XYZ in another language. (Here XYZ stands for a specific section name mentioned below, such as “Acknowledgements”, “Dedications”, “Endorsements”, or “History”.) To “Preserve the Title” of such a section when you modify the Document means that it remains a section “Entitled XYZ” according to this definition.

The Document may include Warranty Disclaimers next to the notice which states that this License applies to the Document. These Warranty Disclaimers are considered to be included by reference in this License, but only as regards disclaiming warranties: any other implication that these Warranty Disclaimers may have is void and has no effect on the meaning of this License.

2. Verbatim Copying

You may copy and distribute the Document in any medium, either commercially or noncommercially, provided that this License, the copyright notices, and the license notice saying this License applies to the Document are reproduced in all copies, and that you add no other conditions whatsoever to those of this License. You may not use technical measures to obstruct or control the reading or further copying of the copies you make or distribute. However, you may accept compensation in exchange for copies. If you distribute a large enough number of copies you must also follow the conditions in section 3.

You may also lend copies, under the same conditions stated above, and you may publicly display copies.

3. Copying in Quantity

If you publish printed copies (or copies in media that commonly have printed covers) of the Document, numbering more than 100, and the Document’s license notice requires Cover Texts, you must enclose the copies in covers that carry, clearly and legibly, all these Cover Texts: Front-Cover Texts on the front cover, and Back-Cover Texts on the

back cover. Both covers must also clearly and legibly identify you as the publisher of these copies. The front cover must present the full title with all words of the title equally prominent and visible. You may add other material on the covers in addition. Copying with changes limited to the covers, as long as they preserve the title of the Document and satisfy these conditions, can be treated as verbatim copying in other respects.

If the required texts for either cover are too voluminous to fit legibly, you should put the first ones listed (as many as fit reasonably) on the actual cover, and continue the rest onto adjacent pages.

If you publish or distribute Opaque copies of the Document numbering more than 100, you must either include a machine-readable Transparent copy along with each Opaque copy, or state in or with each Opaque copy a computer-network location from which the general network-using public has access to download using public-standard network protocols a complete Transparent copy of the Document, free of added material. If you use the latter option, you must take reasonably prudent steps, when you begin distribution of Opaque copies in quantity, to ensure that this Transparent copy will remain thus accessible at the stated location until at least one year after the last time you distribute an Opaque copy (directly or through your agents or retailers) of that edition to the public.

It is requested, but not required, that you contact the authors of the Document well before redistributing any large number of copies, to give them a chance to provide you with an updated version of the Document.

4. Modifications

you may copy and distribute a Modified Version of the Document under the conditions of sections 2 and 3 above, provided that you release the Modified Version under precisely this License, with the Modified Version filling the role of the Document, thus licensing distribution and modification of the Modified Version to whoever possesses a copy of it. In addition, you must do these things in the Modified Version:

- A.** Use in the Title Page (and on the covers, if any) a title distinct from that of the Document, and from those of previous versions (which should, if there were any, be listed in the History section of the Document). You may use the same title as a previous version if the original publisher of that version gives permission.
- B.** List on the Title Page, as authors, one or more persons or entities responsible for authorship of the modifications in the Modified Version, together with at least five of the principal authors of the Document (all of its principal authors, if it has fewer than five), unless they release you from this requirement.

- C. State on the Title page the name of the publisher of the Modified Version, as the publisher.
- D. Preserve all the copyright notices of the Document.
- E. Add an appropriate copyright notice for your modifications adjacent to the other copyright notices.
- F. Include, immediately after the copyright notices, a license notice giving the public permission to use the Modified Version under the terms of this License, in the form shown in the Addendum below.
- G. Preserve in that license notice the full lists of Invariant Sections and required Cover Texts given in the Document's license notice.
- H. Include an unaltered copy of this License.
- I. Preserve the section Entitled "History", Preserve its Title, and add to it an item stating at least the title, year, new authors, and publisher of the Modified Version as given on the Title Page. If there is no section Entitled "History" in the Document, create one stating the title, year, authors, and publisher of the Document as given on its Title Page, then add an item describing the Modified Version as stated in the previous sentence.
- J. Preserve the network location, if any, given in the Document for public access to a Transparent copy of the Document, and likewise the network locations given in the Document for previous versions it was based on. These may be placed in the "History" section. You may omit a network location for a work that was published at least four years before the Document itself, or if the original publisher of the version it refers to gives permission.
- K. For any section Entitled "Acknowledgements" or "Dedications", Preserve the Title of the section, and preserve in the section all the substance and tone of each of the contributor acknowledgements and/or dedications given therein.
- L. Preserve all the Invariant Sections of the Document, unaltered in their text and in their titles. Section numbers or the equivalent are not considered part of the section titles.
- M. Delete any section Entitled "Endorsements". Such a section may not be included in the Modified Version.
- N. Do not retitle any existing section to be Entitled "Endorsements" or to conflict in title with any Invariant Section.
- O. Preserve any Warranty Disclaimers.

If the Modified Version includes new front-matter sections or appendices that qualify as Secondary Sections and contain no material copied from the Document, you may at your option designate some or all of these sections as invariant. To do this, add their titles to the list of Invariant Sections in the Modified Version's license notice. These titles must be distinct from any other section titles.

You may add a section Entitled "Endorsements", provided it contains nothing but endorsements of your Modified Version by various parties — for example, statements of peer review or that the text has been approved by an organization as the authoritative definition of a standard.

You may add a passage of up to five words as a Front-Cover Text, and a passage of up to 25 words as a Back-Cover Text, to the end of the list of Cover Texts in the Modified Version. Only one passage of Front-Cover Text and one of Back-Cover Text may be added by (or through arrangements made by) any one entity. If the Document already includes a cover text for the same cover, previously added by you or by arrangement made by the same entity you are acting on behalf of, you may not add another; but you may replace the old one, on explicit permission from the previous publisher that added the old one.

The author(s) and publisher(s) of the Document do not by this License give permission to use their names for publicity for or to assert or imply endorsement of any Modified Version.

5. Combining Documents

You may combine the Document with other documents released under this License, under the terms defined in section 4 above for modified versions, provided that you include in the combination all of the Invariant Sections of all of the original documents, unmodified, and list them all as Invariant Sections of your combined work in its license notice, and that you preserve all their Warranty Disclaimers.

The combined work need only contain one copy of this License, and multiple identical Invariant Sections may be replaced with a single copy. If there are multiple Invariant Sections with the same name but different contents, make the title of each such section unique by adding at the end of it, in parentheses, the name of the original author or publisher of that section if known, or else a unique number. Make the same adjustment to the section titles in the list of Invariant Sections in the license notice of the combined work.

In the combination, you must combine any sections Entitled "History" in the various original documents, forming one section Entitled "History"; likewise combine any sections Entitled "Acknowledgements", and any sections Entitled "Dedications". You must delete all sections Entitled "Endorsements."

6. Collections of Documents

You may make a collection consisting of the Document and other documents released under this License, and replace the individual copies of this License in the various documents with a single copy that is included in the collection, provided that you follow the rules of this License for verbatim copying of each of the documents in all other respects.

You may extract a single document from such a collection, and distribute it individually under this License, provided you insert a copy of this License into the extracted document, and follow this License in all other respects regarding verbatim copying of that document.

7. Aggregation with Independent Works

A compilation of the Document or its derivatives with other separate and independent documents or works, in or on a volume of a storage or distribution medium, is called an “aggregate” if the copyright resulting from the compilation is not used to limit the legal rights of the compilation’s users beyond what the individual works permit. When the Document is included in an aggregate, this License does not apply to the other works in the aggregate which are not themselves derivative works of the Document.

If the Cover Text requirement of section 3 is applicable to these copies of the Document, then if the Document is less than one half of the entire aggregate, the Document’s Cover Texts may be placed on covers that bracket the Document within the aggregate, or the electronic equivalent of covers if the Document is in electronic form. Otherwise they must appear on printed covers that bracket the whole aggregate.

8. Translation

Translation is considered a kind of modification, so you may distribute translations of the Document under the terms of section 4. Replacing Invariant Sections with translations requires special permission from their copyright holders, but you may include translations of some or all Invariant Sections in addition to the original versions of these Invariant Sections. You may include a translation of this License, and all the license notices in the Document, and any Warranty Disclaimers, provided that you also include the original English version of this License and the original versions of those notices and disclaimers. In case of a disagreement between the translation and the original version of this License or a notice or disclaimer, the original version will prevail.

If a section in the Document is Entitled “Acknowledgements”, “Dedications”, or “History”, the requirement (section 4) to Preserve its Title (section 1) will typically require changing the actual title.

9. Termination

You may not copy, modify, sublicense, or distribute the Document except as expressly provided for under this License. Any other attempt to copy, modify, sublicense or distribute the Document is void, and will automatically terminate your rights under this License. However, parties who have received copies, or rights, from you under this License will not have their licenses terminated so long as such parties remain in full compliance.

10. Future Revisions of this License

The Free Software Foundation may publish new, revised versions of the GNU Free Documentation License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns. See <http://www.gnu.org/copyleft/>.

Each version of the License is given a distinguishing version number. If the Document specifies that a particular numbered version of this License “or any later version” applies to it, you have the option of following the terms and conditions either of that specified version or of any later version that has been published (not as a draft) by the Free Software Foundation. If the Document does not specify a version number of this License, you may choose any version ever published (not as a draft) by the Free Software Foundation.

ADDENDUM: How to use this License for your documents

To use this License in a document you have written, include a copy of the License in the document and put the following copyright and license notices just after the title page:

Copyright © YEAR YOUR NAME. Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.2 or any later version published by the Free

Software Foundation; with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts. A copy of the license is included in the section entitled “GNU Free Documentation License”.

If you have Invariant Sections, Front-Cover Texts and Back-Cover Texts, replace the “with... Texts.” line with this:

with the Invariant Sections being LIST THEIR TITLES, with the Front-Cover Texts being LIST, and with the Back-Cover Texts being LIST.

If you have Invariant Sections without Cover Texts, or some other combination of the three, merge those two alternatives to suit the situation.

If your document contains nontrivial examples of program code, we recommend releasing these examples in parallel under your choice of free software license, such as the GNU General Public License, to permit their use in free software.

Bibliography

- [1] Onn Shehory and Sarit Kraus. Methods for task allocation via agent coalition formation. *Artificial Intelligence*, 101(1-2):165–200, May 1998.