Categorisation as Topographic Mapping between Uncorrelated Spaces

T. Mark Ellison

IO Institute of Linguistics, Australia mark@iolinguistics.com

Abstract. In this paper, I propose a neurophysiologically plausible account for the evolution of arbitrary, categorical mental relationships. Topographic, or structure-preserving, mappings are widespread within animal brains. If they can be shown to generate behaviours in simulation, it is plausible that they are responsible for them in vivo. One behaviour has puzzled philosophers, psychologists and linguists alike: the categorical nature of language and its arbitrary associations between categories of form and meaning. I show here that arbitrary categorical relationships can arise when a topographic mapping is developed between continuous, but uncorrelated activation spaces. This is shown first by simulation, then identified in humans with synaesthesia. The independence of form and meaning as sensory or conceptual spaces automatically results in a categorial structure being imposed on each, as our brains attempt to link the spaces with topographic maps. This result suggests a neurophysiologically plausible explanation of categorisation in language.

1 Introduction

Solomonoff's 1964 paper A Formal Theory of Inductive Inference [29] was a landmark in the study of learning. It brought together the philosophical problem of induction, probability theory, information theory, computability and formal language theory. It related shortness of coding length with goodness of fit, presaging the well-known induction methods of MML/MDL [35, 24]. Of particular interest to cognitive science and natural language processing was his exploration of grammar induction, an analogue of which occurs in the development of every speaker of every natural language. In the current paper, I explore an aspect of language learning which complements the grammar induction task addressed by Solomonoff, asking: how do our brains construct categorial symbols out of the continous inputs they receive? Where the universal Turing machine was the abstract computational notion at the basis of Solomonoff's exploration of learning, the topographic mapping is the basis of the current paper.

Topographic mappings are relations between two spaces which preserve similarity: similar points in one space are related to similar points in the other. As the name suggests, this is an ideal property for maps of territory: nearby geographic locations should be nearby on the map, places far from each other should appear distant on the map.

Topographic maps in the brain were first noticed as a result of the work of Penfield and Boldrey [20, 26]. Directly stimulating the brains of locally anaesthetised epilepsy patients allowed them to explore the qualia associated with particular points on the brain. For the most part, they found that sensations attributed to nearby points on the human surface or entrails resulted from stimulating nearby parts of the brain.

Since this initial work, researchers have both found new topographic structures in the brain [28], and identified previously known structures as resulting from constructing maximally topographic connections from complex inputs [7]. Section 2 discusses these natural topographic mappings in more detail, supporting the premiss that such mappings are widespread and a plausible neurophysiological explanation for cognitive phenomena.

It is possible to define formally what constitutes a topographic map [8], and to construct them algorithmically. Section 3 considers some general definitions and implementations. The crux of the paper comes in section 3.5. When a topographic map is constructed between two spaces on the basis of uncorrelated pairs of inputs and outputs, it is found to exhibit a remarkable form. It shows sectioning of the input into contiguous chunks that map onto a single value or a small connected range of values.

This result is important, because it suggests a neurophysiological explanation for categorisation in language. Language is inherently categorical. The word *cat* consists of the same sequence of phonemes whether said by a small child with a lisp, or sung by an operatic contralto. These variations do not matter for the identity of the word, but there is a limit. Pronounced like *cute*, and the word *cat* is no longer there at all, even though the sounds are similar.

Likewise the meanings of words are categorical. The class refered to by *cat* does not blend, even in response to variations in pronunciation, into *kite* or *coat*. Why is this the case? It is logically possible to imagine a language which was directly iconic: a language of drawn pictures, or as Bickerton imagined, dolphins referring to objects by mimicking their sonar profile [3].

In section 4, I discuss linguistic and cognitive categories and argue that these can be driven by forcing topographic mappings between uncorrelated spaces, one of which consequently becomes the space of representation, and the other, the space of meaning. For example, the well-studied colour categories could arise by mappings from the sensory colour space (along with other sensory and conceptual spaces) onto a space of phonetic expressions.

In fact, we can find direct evidence for this type of categorisation in the experience of synaesthetes. Section 5 relates the categorical nature of synaesthetic associations to a likely neurophysiological explanation: a topographic mapping between cortical areas dealing with the associated qualia.

The final section 6 draws the argument together to support a neurophysiologically plausible account of the development of linguistic and other categorical cognitive processes: topographic mappings created between spaces receiving uncorrelated inputs.

2 Topographic Mappings in the Brain

Topographic mappings relate similar or nearby inputs to similar or nearby outputs. These kinds of mappings are found between the senses and the cortex, between different parts of the cortex (and other parts of the brain), and between the cortex and the motor system. They seem to be a ubiquitous part of neural processing [32].

2.1 Easy to See Mappings

In animal sensory systems, the inputs are primary sensors arranged over a sense field as retinas in the case of sight, the skin in the case of touch, the organs of Corti in the case of hearing, and so on. The outputs are areas of the cortex devoted to sythesising information from these inputs, such as the primary visual cortex. The mapping of the sense of touch onto the cortex is relatively simple.

Penfield and Boldrey [20, 26] sought to express only the area of cortex assigned to particular somatosensory organs in their first homunculus. Consequently, the mapping onto the cortical slice was not shown, but it was implied. If an area of cortex could be mapped onto the hand, so that its appropriate size could be determined, this implied that points within the hand were within a single region in the cortex, and stimuli from outside the hand were processed further away. A more direct mapping between body part and cortex is seen in the homunculus of Penfield and Rasmussen [21]. The topographic nature is clearly visible, for example, a point in the arm is mapped closer to nearer points on the same appendage than to further points. (See [26] for a history of the homunculus.)

Recent fine-grained studies such as [11] have used techniques such as magnetic resonance imaging to create fine-grained maps of the motor and somatosensory cortices. These studies show that while there some overlap of cortical areas responding to the actions of individual fingers, there are orderly somatotopies.

Topographic mappings are also found repeatedly in the visual cortex. The next section deals with the relation between topographic mappings and ocular dominance stripes. Suffice it to say for the moment that the primary visual cortex maps inputs from the visual field according to their position in retina. Recently, specialised topographic mappings of visual inputs that also integrate infromation from other senses have been found [28, 25, 30]. In fact, some [9] assume retinotopicity to determine boundaries for optical representations in the cortex.

These kinds of mapping also occur with auditory stimuli. [15] finds topographic mappings of modulation density by frequency in the cortex.

2.2 Continuous and Discrete - Ocular Dominance Stripes

Some topographic mappings into cortical structure are easy to see. A twodimensional sensory field is mapped into a two-dimensional cortical area, with

geometry preserved. Other kinds of topographic mapping are not so transparent, as they appear to display rampant discontinuity.

Ocular dominance stripes (or columns) were first named such by Hubel and Wiesel [13], although they had noted the phenomenon earlier [12]. They identified columns of cells in the cortex which responded to input from one eye only. These columns were arranged into contiguous striping patterns over the visual cortex.

In work for which they were awarded the Nobel prize, they showed that these stripes were developmentally conditioned and affected by environmental stimuli. When young kittens lost stimuli from one eye, the other took over the cortical areas normally used for signals from the occluded eye. Similar effects have been shown in primates [16].

Ocular dominance stripes have been found in non-mammals such as owls [22] and (in extraordinary circumstances) frogs [14].

Despite our considerable understanding of forces that can cause or disrupt the development of ocular dominance columns, there is no consensus as to their function [1]. There is a model of their development, however, that sees them as a natural (but not necessarily inevitable) consequence of constructing a topographic mapping from two positively correlated planar sources of information onto the cortex.

Goodhill [7] introduced a model for the mapping from eye to brain which is capable of developing both the topographic structure of the mapping as well as the ocular dominance striping. The striping occurs because the inputs to the two eyes are positively correlated in binocular vision. The model predicts that increasing correlation leads to narrower stripes. More to the point, it suggests that when correlations of this kind are present we should not be surprised to find dominance striping along with it.

This prediction seems to have been born out in two results in different species. In one study of frogs [14], the authors implanted a node which developed into a third eye in the growing frog, on the same side as one of the existing eyes, creating binocular vision on that side. The two eyes developed, with some of the visual input from one eye sharing cortical projection with the other. In the shared cortex, ocular dominance stripes resulted.

The two eyes on this side had highly correlated inputs, having overlapping visual fields. The fact that striping occurred in just this case supports Good-hill's analysis of striping as a side-effect of constructing a topographic map from multiple correlated inputs.

Further evidence for this analysis comes from studies of platypus bill sensing reported by Pettigrew [22]. The platypus has electrosensors and mechanosensors intermingled in its bill, and despite these sensors having separate pathways to the brain, they map cortically with striping similar to ocular dominance striping. This is no surprise to the topographic explanation: if the animal is integrating information from two highly correlated topographic arrays (two eyes looking in the same direction or electro- and mechano-sensors spread out on the same surface), then dominance striping is a likely way to bring relevant information together in the brain. So ocular dominance stripes despite, or even because of, their discontinuities can be regarded as optimal topographic mappings from one space, made up of two disjointed but correlated parts, to contiguous space. As we shall see in section 3.5, this is not the only way to get chunking and discontinuities from topographic mappings.

In summary then, topographic mappings arise frequently in the brain; the more we look for them, the more we find them. They integrate knowledge from different receptive fields, and even different sense types. Furthermore, they can give rise to complex discontinuous structures.

3 The Topographic Extrapolation

This section presents a measure of topographicity, and extends it to finding the most topographic extensions of given data, in a manner similar to that described by Ellison [6].

3.1 Measuring Topographicity

Goodhill and Sejnowski [8] offered the general evaluation measure for the topographicity of functions shown in (1) (symbols have been changed to avoid confusion).

$$C(f) = \sum_{i \in I} \sum_{j \in I} s_I(i, j) s_O(f(i), f(j))$$
(1)

Here s_I and s_O are similarity measures on the input and output spaces respectively. C(f) is the measure of how topographic the space is, with larger values for more topographic mappings. Both the s_I and s_O are assumed to be everywhere non-zero. Consequently C(f) (if finite) is maximised for f which match small values of s_I with small values of s_O and large with large. See Goodhill and Sejnowski's discussion for more details.

The definition can be extended to apply to any finite collection D of inputoutput data points (2).

$$C(D) = \sum_{(i,o)\in D} \sum_{(j,p)\in D} s_I(i,j) s_O(o,p)$$
(2)

3.2 Extrapolation

Solomonoff [29], following Carnap [5], couched learning problems as sequence extrapolation coding problems. He started by defining a cost for a sequence as the length of the shortest program for a given universal Turing machine which would generate that sequence. From this was constructed an answer to the extrapolation problem: the best extrapolation of a sequence is the one which gives the smallest increment in the cost function.

The same approach can determine the most topographic extrapolation from a given set of input-output pairs D. Since s_I and s_O are similarity measures, rather than cost functions or distances, larger C(D) is better. Consequently, we seek to maximise the change in C(D) resulting from the addition of the new pair (x, y), as expressed in (3).

$$C(D \cup \{(x,y)\}) - C(D) = 2 * \sum_{(i,o) \in D} s_I(i,x) s_O(o,y) + s_I(x,x) s_O(y,y) \quad (3)$$

We assume that the self-similarity of each point in the input space and the output space are constant: $s_I(x, x)$ does not depend on x, nor $s_O(y, y)$ on y. This means that the whole equation is maximised when the summation is maximised, allowing us the concise form shown in (4) for the best extrapolation.

In what follows, we will write S(x, y|D) for the term $\sum_{(i,o)\in D} s_I(i, x) s_O(o, y)$.

$$E(D) = (x, y) | x \in I, y \in O \text{ that minimise } S(x, y|D)$$
(4)

The extrapolation can be limited to selecting the best output for a given input x. If this is done for all values of $x \in I$, we excapalate a most topographic function fitting the data (5).

$$F(D) = \{(x, y) | x \in I, y \in O \text{ where } y \text{ uniquely minimises } S(x, y|D)\}$$
(5)

Frequently, F(D) is only a partial function. If there are multiple maxima for S(x, y) for some x, then no y uniquely minimises this function. For these values of x, the function remains undefined.

3.3 A Normal Similarity Measure

In the simulations which follow, I use a probability-based similarity measure. This measure parallels the notion of confusion probabilities as discussed by the author in an earlier paper on induction [6]. It is defined, for an arbitrary space in equation (6). One parameter to the measure is scale constant σ assumed constant over the whole space.

$$s(x,y) = N(x-y;\sigma) \tag{6}$$

The measure has not been normalised over either x or y to retain its commutativity, so will not act here as a probability distribution. It can be proven that given this similarity measure, for any finite data set D, the extrapolated function F(D) is continuous wherever it is defined.

Extrapolation from Highly Topographic Functions $\mathbf{3.4}$

At this point, it is worth looking at the extrapolation from data which is highly topographic. For simplicity, take the one-dimensional unit interval [0,1]as both input and output space.

A program, dubbed OPTOPO, was developed in PYTHON to extrapolate pointwise the most topographic function matching a given input data set, based on the optimisation derived in section 3.2.

Data was generated by mapping each of 20 points equally spaced along the unit interval to itself. These were then used to extrapolate values covering the same interval at 0.001 spacing. A graph of the resulting Fig. 1: Topographic



extrapolation extrapolations appears in figure 1. The simi- from 20 points of io = i on the unit larity measures are normal density functions interval to the 1000 points.

as described in section 3.3, with scale constants σ set for input and output spaces to 0.04.

In general, we can expect strongly topographic data, in the sense of equation (2), to give rise to extrapolations matching those functions.

$\mathbf{3.5}$ **Independently Varying Spaces**

In the last example, a thousand points were extrapolated from 20 belonging to the already topographic identity function. This function was recovered, apart from small curvature at the extremes ends of the interval. Now we ex-



[0,1].

plore the other extreme: what happens when the training data is random and uncorrelated?

Training data was generated, consisting of 500 points with random input and output values drawn from a flat distribution over the unit interval. A scatter plot of this data is shown in figure 2.

Applied to this training data, OPTOPO re-Fig. 2: Input data to the topo- turned the function shown in figure 3 as graphic mapping between uncorre- the most topographic that could be deduced lated spaces: 500 points in $[0,1] \times$ from the data. The most remarkable feature of this graph is that it shows a largely

smooth, continuous function with only a handful of discontinuities. Between the discontinuities, we have continuous maps from input to output. Although the mapping output is usually fairly level, there is sometimes variation. In these cases the variation is smooth. An input point between two close neighbours maps onto an output point between the projections of those neighbours.

The explanation of this combination of smooth mapping plus discontinuity

lies in the fact that we are looking for the most topographic output value o for each input i. The smooth similarity measures turn the scattered host of input

data points into a continuously contoured scalar field over the input-output plane. Constructing the most topographic extrapolation function is a process of maximising the evaluation for each input value. The scalar field over the plane has ridges of higher values, and these are what we see defining the contiguous categories in the input. Transitions from one ridge to another happen when a falling ridge of values is overtaken by a ris- Fig. 3: The most topographic (i.e. ing ridge, as the focus moves along input val- the most probable) function account-



ues. This transitional discontinuity is of the ing for the uncorrelated input data. kind made famous in the Catastrophe Theory of Rene Thom [33].

Functions with this kind of discontinuity define a categorisation of the input. Inputs within a connected region belong to one category. Inputs from regions separated by discontinuities come from different categories. Topographic mappings offer an easy route from continous spaces to categorisation, and this suggests elegant new explanations for the origin and nature of language.

4 Explaining the Categorical Nature of Language

Language is both symbolic and categorical. It is symbolic because there is no intrinsic, or to use Peirce's [19] terms, no indexical and no iconic, relation between the space of meanings, and the space of linguistic representations. No in-



Fig. 4: Iconic representation of fire.

dexical relationship means that there is no physical association of our units of meanings with what they represent: our word fire does not share the same relationship with actual fire that actual smoke does. No iconic relationship means that there is no structure-preserving mapping which relates the word *fire* to what it represents. In contrast, the sign for fire shown in figure 4, is an iconic representation: while this image has no physical connection to fire, the qualia of looking at the image share structural similarities to a view of a fire. But not being iconic or indexical

does not account for the categorical nature of natural language symbols. So language is a connection between uncorrelated spaces of meaning and representation. Could language be a topographic mapping between these spaces? If so, it would explain the categorical nature of both meanings (eg. colour classes, genders, emotions, etc.) and forms in language (eg. the phonological forms of words). If this were the case, we would expect two features of language: arbitrary categorisation, with categories varying from language to language, and a globally non-topographic function from form to meaning.

Language categories are arbitrary, and they certainly vary from language to language, although there may be functional pressures which limit variation.

Variation in the number and phonetic realisation of phonemes is one example. Language-specificity is more hotly debated, however, with regard to semantic categories. Nevertheless researchers report tracking shifts in categorisation as children adapt to the language they are acquiring [4, 17, 18].

Vocabulary also shows no strong large scale topographic relation between meaning and form. Shillcock et al [27] and Tamariz [31] looked for large-scale systematicity in the meanings of monomorphemic words in English and Spanish respectively. Effects were found, but primarily for localised domains such as pause fillers, and for limited phonetic features.

So the evidence agrees with a model in which word forms are related to their meanings by a topographic mapping linking uncorrelated spaces.

5 Synaesthesia

Synaesthesia is the leakage of qualia from one sensory input to another. For example, someone might see – in their mind's eye – the colour *red* when they hear the word *hurry*, or smell olives when seeing square objects. For an overview of synaesthesia, see [10].

De Thornley Head [34] finds that pitch-colour synaesthetes have reliable mappings from pitches onto colours (unlike non-synaesthetes), and takes this as evidence that synaesthesia is a perceptual phenomenon apart from memory, metaphor or imagery. He interprets the mapping itself in the following way. Pitches and colours are broken into regions by some unclear mechanism. These regions are then associated in an arbitrary fashion. There can be smooth topographic mappings within the regions so associated, but no such connections hold when region boundaries are crossed.

These findings agree with a model in which synaesthesia is the result of additional connectivity, a topographic map, linking an auditory representation with a representation of colour. That synaesthesia results from structural brain differences is supported experimentally [2, 23]. For the most part, auditory and colour stimulus arrive from the senses uncorrelated. Consequently, the topographic map developed has the form of figure 3: discontinuous regions of input mapped smoothly onto arbitrary and disconnected regions of output.

6 Conclusion

The purpose of this paper was to propose a neurophysiologically plausible explanation for the formation of arbitrary categorisations. The core of this explanation is that, as seen in section 2, topographic mappings are commonplace in the brain, turning up even when their nature is less than obvious, as in the case of ocular dominance striping. To make the explanation concrete, a formal description of topographic mappings was presented in section 3. This section finished with the core result of this paper: the most topographic linkage between two uncorrelated spaces can be a categorical mapping with, at most, in-category smooth mapping. Section 4 and section 5 explored some of the phenomena which

might be explained using this result. Monomorphemic vocabulary shows little large-scale correlation of similarity of phonetic form and similarity of meaning, instead linking somewhat arbitrary categories of meaning to arbitrary (though systematic) categories of form. Similarly, pitch-colour synaesthetics show only local smoothness in their mapping of one modality to the other; the mapping consists of connections between somewhat arbitrary chunks of the pitch and colour spaces. Topographic mappings between uncorrelated spaces explain both these phenomena.

References

- Adams, D.L., Horton, J.C.: Ocular dominance columns: Enigmas and challenges. The Neuroscientist 15(1), 62 –77 (Feb 2009)
- Bargary, G., Mitchell, K.J.: Synaesthesia and cortical connectivity. Trends in Neurosciences 31(7), 335–342 (Jul 2008)
- 3. Bickerton, D.: King of the sea. Random House, 1st edn. (1979)
- Borovsky, A., Elman, J.: Language input and semantic categories: A relation between cognition and early word learning. Journal of Child Language 33(04), 759– 790 (2006)
- Carnap, R.: Logical Foundations of Probability. University of Chicago Press, Chicago, IL (1950)
- Ellison, T.M.: Induction and inherent similarity. In: Similarity and Categorization, pp. 29–50. Oxford University Press (2001)
- 7. Goodhill, G.J.: Topography and ocular dominance: a model exploring positive correlations. Biological Cybernetics 69(2), 109–118 (1993), PMID: 8373882
- Goodhill, G.J., Sejnowski, T.J.: A unifying objective function for topographic mappings. Neural Computation 9(6), 1291–1303 (Jun 1997)
- Greenlee, M.W.: Human cortical areas underlying the perception of optic flow: brain imaging studies. International Review of Neurobiology 44, 269–292 (2000), PMID: 10605650
- Harrison, J.E., Baron-Cohen, S.: Synaesthesia: Classic and Contemporary Readings. Wiley-Blackwell (Jan 1997)
- Hlutk, P., Solodkin, A., Gullapalli, R.P., Noll, D.C., Small, S.L.: Somatotopy in human primary motor and somatosensory hand representations revisited. Cerebral Cortex 11(4), 312 –321 (Apr 2001)
- Hubel, D.H., Wiesel, T.N.: Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. The Journal of Physiology 160(1), 106–154.2 (Jan 1962), PMID: 14449617 PMCID: 1359523
- Hubel, D.H., Wiesel, T.N.: Receptive fields and functional architecture of monkey striate cortex. The Journal of Physiology 195(1), 215 –243 (Mar 1968)
- Ide, C., Fraser, S., Meyer, R.: Eye dominance columns from an isogenic doublenasal frog eye. Science 221(4607), 293 –295 (Jul 1983)
- Langers, D.R.M., Backes, W.H., van Dijk, P.: Spectrotemporal features of the auditory cortex: the activation in response to dynamic ripples. NeuroImage 20(1), 265–275 (Sep 2003), PMID: 14527587
- Le Vay, S., Wiesel, T.N., Hubel, D.H.: The development of ocular dominance columns in normal and visually deprived monkeys. The Journal of Comparative Neurology 191(1), 1–51 (May 1980)

- Majid, A., Boster, J.S., Bowerman, M.: The cross-linguistic categorization of everyday events: A study of cutting and breaking. Cognition 109(2), 235–250 (Nov 2008)
- Majid, A., Gullberg, M., Staden, M.v., Bowerman, M.: How similar are semantic categories in closely related languages? a comparison of cutting and breaking in four germanic languages. Cognitive Linguistics 18(2), 179–194 (Sep 2007)
- Peirce, C.S., Hartshorne, C., Weiss, P., Burks, A.W.: Collected Papers of Charles Sanders Peirce. Harvard University Press (1931-1958), http://books.google.com/ books?id=G71zSoUFx1YC
- 20. Penfield, W., Boldrey, E.: Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation. Brain 60(4), 389–443 (Dec 1937)
- 21. Penfield, W., Rasmussen, T.: The Cerebral Cortex of Man. Macmillian (1950)
- Pettigrew, J.D.: Bi-sensory, striped representations: comparative insights from owl and platypus. Journal of Physiology-Paris 98(1-3), 113–124 (2004)
- Ramachandran, V., Hubbard, E.: Synaesthesia a window into perception, thought and language. Journal of Consciousness Studies 8(12), 3–34 (2001)
- 24. Rissanen, J.: Modeling by shortest data description. Automatica 14, 465–471 (1978)
- Schluppeck, D., Glimcher, P., Heeger, D.J.: Topographic organization for delayed saccades in human posterior parietal cortex. Journal of Neurophysiology 94(2), 1372–1384 (Aug 2005), PMID: 15817644
- Schott, G.D.: Penfield's homunculus: a note on cerebral cartography. Journal of Neurology, Neurosurgery, and Psychiatry 56(4), 329–333 (Apr 1993), PMID: 8482950 PMCID: 1014945
- 27. Shillcock, R., Kirby, S., McDonald, S., Brew, C.: Filled pauses and their status in the mental lexicon (2001)
- Silver, M.A., Kastner, S.: Topographic maps in human frontal and parietal cortex. Trends in Cognitive Sciences 13(11), 488–495 (Nov 2009)
- Solomonoff, R.J.: A formal theory of inductive inference. Information and Control pp. 1–22, 224–254 (1964)
- Swisher, J.D., Halko, M.A., Merabet, L.B., McMains, S.A., Somers, D.C.: Visual topography of human intraparietal sulcus. The Journal of Neuroscience 27(20), 5326 –5337 (May 2007)
- Tamariz, M.: Exploring systematicity between phonological and contextcooccurrence representations of the mental lexicon. The Mental Lexicon 3, 259–278 (Sep 2008)
- Thivierge, J., Marcus, G.F.: The topographic brain: from neural connectivity to cognition. Trends in Neurosciences 30(6), 251–259 (Jun 2007)
- Thom, R.: Structural stability and morphogenesis. Addison Wesley Publishing Company (1989)
- de Thornley Head, P.: Synaesthesia: Pitch-Colour isomorphism in RGB-Space? Cortex 42(2), 164–174 (2006)
- Wallace, C.S., Boulton, D.M.: An information measure for classification. Computer Journal 11(2), 195–209 (1968)