A CONTEXT-SENSITIVE META-CLASSIFIER FOR COLOR-NAMING

by

Rony Daniel Kubat

B.S. Computer Science, MIT 2001
B.S. Mechanical Engineering, MIT 2001

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Master of Science in Computer Science and Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2008

© Rony Daniel Kubat, MMVIII. All rights reserved.

The author hereby grants to MIT permission to reproduce and distribute publicly paper and electronic copies of this thesis document in whole or in part.

Author ....
Department of Electrical Engineering and Computer Science
October 1, 2007

Certified by ....
Deb K. Roy
Associate Professor
Thesis Supervisor

Accepted by ....
Terry P. Orlando
Chairman, Department Committee on Graduate Students

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
APR 0 7 2008
LIBRARIES

ARCHIVES

A Context-Sensitive Meta-Classifier for Color-Naming

by

Rony Daniel Kubat

Submitted to the Department of Electrical Engineering and Computer Science
on October 1, 2007, in partial fulfillment of the
requirements for the degree of
Master of Science in Computer Science and Engineering

Abstract

Humans are sensitive to situational and semantic context when applying labels to colors. This is especially challenging for algorithms which attempt to replicate human categorization for communicative tasks. Additionally, mismatched color models between dialog partners can lead to a back-and-forth negotiation of terms to find common ground. This thesis presents a color-classification algorithm that takes advantage of a dialog-like interaction model to provide fast-adaptation for a specific exchange. The model learned in each exchange is then integrated into the system as a whole. This algorithm is an incremental meta-learner, leveraging a generic online-learner and adding context-sensitivity. A human study is presented, assessing the extent of semantic contextual effects on color naming. An evaluation of the algorithm based on the corpus gathered in this experiment is then tendered.

Thesis Supervisor: Deb K. Roy
Title: Associate Professor

Acknowledgments

I have been granted a great opportunity to work among such a talented and interested cadre as I have found at Cognitive Machines, the research group where I've spent the last two years. First, I would like to thank my intrepid advisor, Deb Roy who has been all I want from a mentor, providing inspiration and guidance but also freedom and trust. I also want to thank the National Science Foundation and the American taxpayer, without whom it would have been impossible for me to find my home at the Media Lab amongst the cogmac'ers. Many thanks to all the members of cogmac, with whom discussions have been informative and inspiring and critiques strengthening. Special thanks are deserved by Dan Mirman who helped tremendously in the design and analysis of the experiment described here. Finally, thank you to my family: Peter, Hadassa and Emilie who have been unyielding sources of support, encouragement and creative inspiration.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

What does it mean to say A is black? Rather it was as if I were discovering colors for the first time: red was quite cheerful, *fire red*, but perhaps too strong. No, maybe yellow was stronger, like a light suddenly switched on and pointed at my eyes. Green made me feel peaceful. The difficulties arose with the other little squares. What's this? Green, I said. But Gratarolo pressed me: what type of green, how is it different from this one? Shrug. Paola explained that this one was emerald green and the other was pea green. Emeralds are gems, I said, and peas are vegetables that you eat. They are round and they come in a long, lumpy pod. But I had never seen either emeralds or peas. Don't worry, Gratarola said, in English they have more than three thousand terms for different colors, yet most people can name eight at best. The average person can recognize the colors of the rainbow: red, orange, yellow, green, blue, indigo, and violet—though people already begin to have trouble with indigo and violet. It takes a lot of experience to learn to distinguish and name the various shades, and a painter is better at it than, say, a taxi driver, who just has to know the colors of traffic lights.

The Mysterious Flame of Queen Loana, Umberto Eco (p. 21)

CHAPTER 1

INTRODUCTION

Anjou Pear. Frolic. Capri. Bagel. Heartthrob. Camelback. Flip through the catalog of paints at your local hardware store, and these are the kinds of names you'll find, each a coding for a specific combination of inks. None of these terms are universally used for the subtle hues of the spectrum. Calling your mother and telling her that you're painting your bedroom "summer day" won't quite convey the off-peach tone. Color, though, is one of the key ways we refer to things in our world. What kind of wine would you like with your sirloin? Describe the car that left the scene of the crime...

Somehow, through multiple layers of perception and cognition, we transform a patch of light hitting our retinas into a label; a color name. And what's more, that name is simultaneously stable to radical shifts in lighting, and malleable to the situation. In the sciences, color has been a window into the mind. By carefully controlling the light striking the light-sensitive cells of the eye, we have learned about neural coding at the lowest levels of perception. By surveying languages of the world, we have discovered universals in the categories of color and hypothesized about what these universals mean for the evolution of language and of thought. Engineers have arrived first from a different standpoint: how can colors be reproduced accurately. How can we represent them compactly? Transmit them? And now, how can we categorize them?

This thesis touches on both the science of the human perception of color and the engineering of distinguishing one color from another, and through this investigation connects with a broader issue of classification with a frame of reference: contextual dependence.

1.1 Context Sensitivity

Politicians complain that their words are quoted out of context; that a phrase, removed from the particulars of situation, takes on meaning mismatched—or worse yet, contradictory—to what was intended. Word meanings are mutable to the context of their use. What's meant of "weight" when comparing a heavy feather to a light bowling-ball? What of discipline when a father speaks to his son or a warden to a prisoner? Any model of word meaning must take context into account, but formulating a general model is an enormous undertaking. Here, I grasp at one narrow manifestation of a context's effect on meaning in the domain of color naming.

1.2 Motivation and Inspiration

The work presented here was initially motivated by a specific application in linguistic grounding, the connection of words to the real world [34]. Trisk is a robot at the MIT Media Lab designed to interact with objects placed on a table before it, and to communicate about them with humans by speech. Trisk visually identifies objects of interest by segmenting camera input based on color.[1] We found this color segmentation fragile to changes in lighting conditions, shadows, and the specular reflections of the objects in view. The work of this thesis began in part as a venture to find a robust color-based method of image segmentation. Trisk uses color terms to refer to objects and can respond to imperatives such as "put the green one to the left of the blue one." Like an art dealer describing a painting, Trisk must match color

[1]Computer vision is not the focus of the work presented in this thesis, though for those interested, a survey of color-based segmentation techniques can be found in [8].

to label. It's in this more direct use of color classification that the system described here will likely find more immediate use.

The name of the context-sensitive classification system I developed is Context Dreaming. The *dreaming* half of the name comes from one of the most direct inspirations for the system. Daoyun Ji and Matthew Wilson of the Picower Center for Learning and Memory at MIT recently published a paper [18] supporting a proposal for memory consolidation. In this paper, they report rats playing back memories while dreaming. Perhaps the kernel of this notion of memory playback during an "off-line" time could be directly implemented by a computer?[2] Thus came the two-phase interrogative learning model that Context Dreaming employs. Humans also appear to learn by two different routes. There is a fast "in-context" system, and a slower learning mechanism which consolidates and integrates multiple experiences [25].

1.3 Straddling Two Worlds

This thesis spans both cognitive science and computer science. The contribution to the cognitive sciences are the results of an experiment I performed to assess semantic contextual influence on color categorization. These results confirm that even abstract context can affect low-level perception, and raise questions about the mechanism that causes this effect. In the computer sciences, I have designed a meta-classification algorithm which takes advantage of a real-world interrogative interaction model and can transform online-learners into context-sensitive online-learners.

1.4 Outline

The next six chapters describe the framework I designed to add context sensitivity to online classifiers. The next chapter begins with a short review of the science of color—its representation and partitioning—and describes

[2] The Context Dreaming algorithm shares the nomenclature of the wake-sleep algorithm for neural-networks[17], but not the mechanics.

the relevant work which frames this thesis. Chapter 3 describes in detail the Context Dreaming algorithm. Next is a report of the experiment I designed and performed to quantify some semantic context effects on color naming. The corpus gathered in that experiment is used to evaluate the Context Dreaming algorithm in Chapter 5. Finally, I conclude with a proposal for future directions for this line of research.

CHAPTER 2

BACKGROUND AND RELATED WORK

The study of how people select names for colors has a rich history. Color can be seen as a window into cognition—a direct route to address at least one aspect of the nature versus nurture debate. Are color labels independent of language and tradition or do upbringing and culture directly shape perception? The goal of this chapter is to briefly introduce the key concepts which frame this thesis, and provide context for the choices I have made. The first part of this chapter discusses the science of color and its perception by humans, especially focusing on representations of color. It is on this substrate that parts of this thesis are built. The second part of the chapter gives a condensed introduction to research on color-categorization, and describes some related work on computational models of color-naming. The final part considers some related work on context sensitivity and meta-classifiers. Those familiar with these topics may skip the sections (or the chapter) entirely, without losing critical information about Context Dreaming, its evaluation, or the context-dependent findings discussed in Chapter 4.

2.1 A Brief Introduction to the Science of Color

Imagine you are sitting at your kitchen table at dusk, a basket of fruit before you. A clear sky outside illuminates the room dimly, while an incandescent lamp overhead casts a pool of light upon the bowl. What happens when you "see" the apple on your desk? Light from the sky and the lamp strike the surface of the apple, where it is filtered and reflected into your eyes. There, the light is absorbed by your retina and translated into signals which travel to your brain. Somehow, you decide the color of the apple is red. I will use this simple example to help introduce some key concepts which will take us from the illuminant to the retina.

Light is a continuous spectrum of electromagnetic energy. The range of the spectrum visible to the human eye are the wavelengths between 300nm and 700nm. Purely spectral light—monochromatic light composed of one particular wavelength—is, in a sense, a pure color. Rainbows are made from these pure colors. Partitioning the visible spectrum into colors, we see violet at 300nm range though deep red at the 700nm. Most sources of light, though, radiate a *distribution* of the spectrum rather than a specific wavelength, a bumpy but continuous spread of energy. The most idealized case is that of a black-body, a material whose spectral radiation is defined only by its temperature and related by Plank's Law:

$$I(\lambda, T) = \frac{2hc^2}{\lambda^5} \frac{1}{e^{\frac{hc}{\lambda kT}} - 1}$$

What we refer to as "white" light is a complex distribution across the spectrum—in fact, there is no single standard for white light. The International Commission on Illumination (known more commonly by the acronym for its French name, the *Commission internationale de l'éclairage*, CIE) has defined a number of standard illuminants approximating common sources of white light. Figure 2.1 shows the spectral distribution of a few of these illuminants, including a black-body source. These standard illuminants are the basis for the *white points* used by the color representation standards discussed below.

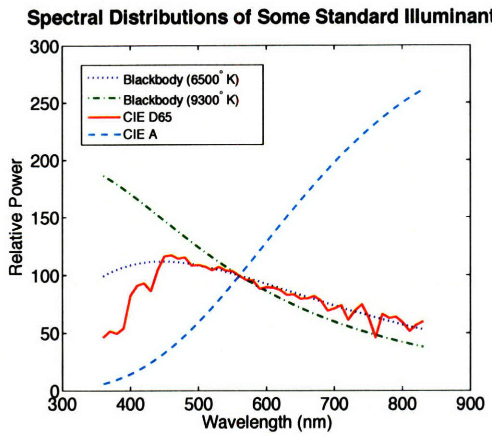


Figure 2.1: The spectral distribution of a few common illuminants. The smooth curve is the idealized black-body radiation of a $6500°K$ source. The bumpy curves are experimentally measured distributions of the CIE standard **D65** (sRGB and television, western Europe daylight) and **A** (incandescent) sources.

Humans compensate for the radical spectral differences of white light, a phenomenon called *color constancy.* Once you've adjusted to the ambient lighting, a sheet of paper *looks* white, whether seen at dusk or at noontime on a sunny day. Color constancy is a *perceptual* effect. Computational color constancy, sometimes called white-balancing, is a long-studied problem. See [1] for a comparison of different algorithms.

The scene in the kitchen has two primary light sources: the sky, which has a bluish tint, and the incandescent bulb with a yellowish tint. We can model these light sources with the CIE illuminants D65 and A, respectively. Light hitting the apple is a superposition of these sources. When this light strikes the surface of the apple, it is selectively absorbed and reflected—transforming the incident spectral distribution into the final one which reaches your eyes.

2.1.1 Biological Basis

Light striking the human retina is absorbed by one of two types of light-sensitive cells. One of these types, *rods*, are sensitive to dim light, but not used to distinguish colors and will not be discussed here further. The color-sensitive type, known as *cone* cells, come in three varieties,[1] each of

[1] Colorblindness is a genetic limitation in which only two types of cones are present. There is some evidence of human tetrachomats (people with four types of cone cells, with four distinct photopigments), but as of this writing, very few have been found.

which uses a distinct photopigment to selectively absorb light spectra. The excitation of a cone cell is a function of both the incoming spectra and the absorption of the cell's photopigment:

$$L(\lambda) = \int I(\lambda)a(\lambda)d\lambda$$

where $L(\lambda)$ is the cell's response, $I(\lambda)$ is the spectral power of the incident light and $a(\lambda)$ is the absorbance of the cone cell's photopigment. The human visual response to color can thus be quantified by the rates of excitation of the three types of cone cells. A consequence of this tristimulus representation of color is metamerism: two distinct color spectra may result in the same responses by the three cone types.

2.1.2 Oppositional Color Theory

The earliest models of color were split into two camps: Isaac Newton leading from a physical substrate based on color spectra, and Johann Goethe from empirical experiments on human perception. Goethe describes his theory in *Theory of Colours*[41]. Introduced in the book is Goethe's color wheel, a symmetric ring where colors are "arranged in a general way according to the natural order, and the arrangement will be found to be directly applicable [...]; for the colours diametrically opposed to each other in this diagram are those which reciprocally evoke each other in the eye."[2]

Goethe's notion that colors are arranged in oppositional pairs anticipated the *opponent process* model proposed by Ewald Hering [16], where colors are encoded as the difference between tristimulus values. As a consequence, red and green oppose each other, as do blue and yellow.

The debate between opponent models and tristimulus models carried though the 1800s, with Hering and Hermann von Helmholtz as prominent proponents of each respective theory. Today's consensus is a combination of both theories, tristimulus tied to low-level perception and opponent colors at higher levels of cognition.

[2] paragraph #50

2.1.3 Color Spaces

The desire to faithfully capture and reproduce color gave rise to the question of how to accurately and efficiently represent color. A *color space* is a method of mathematically encoding color. The *gamut* of a color space is the set of colors representable in that space. Discussed here are a few of the prominent color spaces used for scientific and color reproduction purposes, all of which are represented as a triple of numbers. Formulae for converting between the color spaces described here can be found in [43].

LMS

The three types of cone cells in the human eye contain photopigments which, at first approximation, absorb light in long, medium and short wavelengths. The LMS (long-medium-short) color space gets its name from this fact. LMS triples represent the excitation of the three types of cone cells and so LMS space is most closely grounded to the physiological response of human vision. Nevertheless, LMS is almost never used in either color capture or reproduction due to mismatches in the sensitivity between the three types, the linearity of their measure, and the difficulty relating LMS values to color reproduction by screen or printing. LMS space is linearly related to XYZ (see below). The LMS gamut spans all visible colors.

XYZ and its variants

In 1931, the International Commission on Illumination (CIE) formulated a standard representation of color named XYZ. XYZ was one of the first scientifically defined color representations and has remained the basis of many of color spaces later developed. Each of the three components of XYZ (which roughly correspond to red, green and blue) are linearly tied to the human LMS tristimulus responses.

The XYZ standard is based on a color matching experiment in which subjects were presented with two patches of color, separated by a screen. On one side was a test color of fixed intensity; on the other, a combination of three monochromatic sources whose brightness could be adjusted. By

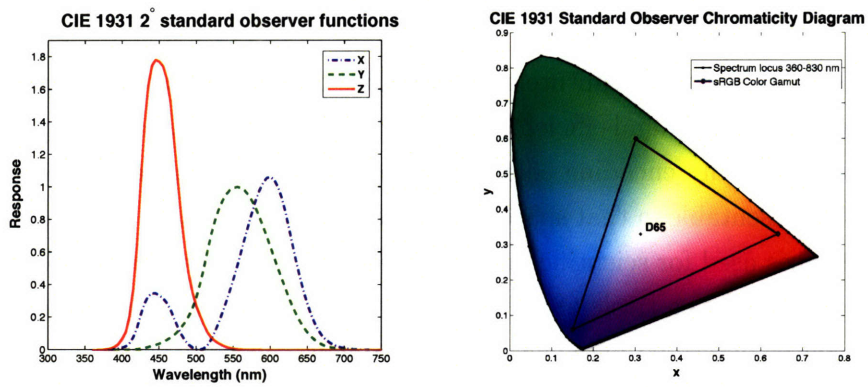


Figure 2.2: The CIE XYZ color matching functions and the xy chromaticity diagram showing the sRGB gamut and white-point.

manipulating the primaries, subjects found a metameric color which could be quantified by the intensities of the three primaries. From this data, CIE created the *standard observer color matching functions* $\bar{x}$, $\bar{y}$ and $\bar{z}$, each of which is a function over wavelength λ. A color in the XYZ color space can then be defined by the equations:

$$X = \int_0^\infty I(\lambda)\bar{x}(\lambda)d\lambda, \qquad Y = \int_0^\infty I(\lambda)\bar{y}(\lambda)d\lambda, \qquad Z = \int_0^\infty I(\lambda)\bar{z}(\lambda)d\lambda$$

Figure 2.2 shows the CIE XYZ color matching functions.

An XYZ triple encodes both the color of light as well as its intensity. A standard decoupling normalizes x and y into a new space xyY defined as:

$$x = \frac{X}{X+Y+Z}, \qquad y = \frac{Y}{X+Y+Z}$$

The two normalized chromaticity coordinates x and y encode color while the third coordinate scales for intensity. The locus of monochromatic light, swept through the spectrum of visible colors traces a horse-shoe shaped arc whose inner area contains all colors visible to humans. Points outside this arc represent ratios of excitation impossible for the three types of cone cells. They are called *imaginary colors*.

RGB and its variant sRGB

The reproduction of color for television and computer displays is by combination of three color primaries of red, green and blue. The brightness of each primary can be represented as a normalized number in the range $[0, 1]$. The chromaticity of the three primary colors forms a triangle which defines the gamut of the RGB space. One other factor completes an RGB space: a white-point. This XYZ triple, corresponding to the "color" of white light, provides a parameter to a transformation which can be used to adapt the RGB space to the color temperature of the viewing environment.

Relevant to this thesis is one particular RGB standard named sRGB, developed by Microsoft and HP to standardize monitors and printers. The chromaticity of the sRGB primaries are based on standard phosphors for CRT displays and the white-point set at CIE D65. Figure 2.2 shows the spectral locus and the gamut and white-point of the sRGB standard.

Perceptually linear color spaces and CIE L*a*b*

A problem with the color spaces described above, especially with regards to color naming, is their perceptual non-linearity. Euclidean distance in XYZ or RGB is not comparable to perceptual distance. For each visible color in the xy chromaticity diagram, there is an ellipse of nearby colors which are perceptually indistinguishable. The size of these *MacAdam ellipses* [23] varies, smallest in the blues and growing larger toward the greens and reds. There have been a number of attempts to define color spaces for which the MacAdam ellipse stays approximately the same size throughout the color space. Moreover, the goal of these color spaces is to make Euclidean distance a parallel measure of perceptual distance gathered experimentally.

CIE L*a*b* was the CIE's 1976 attempt to define a color space that balanced perceptual linearity with straightforward conversion to and from XYZ. The three components of a CIE L*a*b* triple are luminance (a measure of brightness) and two color-difference chromaticity values a* and b*, which roughly encode differences between green and magenta, and blue and yellow, respectively. In that sense, though a triple, CIE L*a*b* can be considered an opponent color space. The $*$ in CIE L*a*b* notes that each component is converted from XYZ with an exponential—better matching human logarithmic response. L* values vary between zero and one hundred; a* and b* values vary in the range $[-128, 128]$.

Although much closer to perceptually linear than XYZ, CIE L*a*b* is not perfect. Other perceptually linear color spaces have been proposed, including the OSA Uniform Color Scales Samples [31], CIE L*u*v*, and NPP [22]. Mojsilovic describes a non-Euclidean distance metric which compensates for irregularities in CIE L*a*b* [29].

2.2 Color Categorization

This section summaries some previous work in color categorization both in the cognitive and computer sciences. In the cognitive sciences especially, color classification has been an active field of research, perhaps because of the ease with which experiments can be created and replicated as well as the close connection between raw stimulus and semantic structure.

2.2.1 ... in the Cognitive Sciences

In 1969, Berlin and Kay published *Basic Color Terms: Their Universality and Evolution*[3], a collection of their research about the naming of colors across cultures and languages. In their key experiment, a standard palette of color chips were named by participants speaking different native languages. Language-specific aggregate mappings from colors to names were collated from this data. Berlin and Kay put forth two hypotheses: that (1) there is a restricted and universal catalog of color categories and (2) languages add these categories in a constrained order. Languages with only two color terms would have terms for black and white. Languages with three have black, white and red. Those languages with four have terms for the former three, plus green or yellow. The hierarchy for color terms proposed was:

$$\begin{array}{c} \text{white} \\ \\ \text{black} \end{array} < \text{red} < \begin{array}{c} \text{green} \\ \\ \text{yellow} \end{array} < \text{blue} < \text{brown} < \begin{array}{c} \text{purple} \\ \text{pink} \\ \\ \text{orange} \\ \text{gray} \end{array}$$

Data collection for this experiment has continued through the World Color Survey (WCS) [19] and a recent analysis of this data argues that the partitioning of color categories follows an optimal partitioning of the space [33], lending strength to the argument that human partitioning of color space into categories is in large part bound to the physiology of human vision. Low-level color perception, though, is influenced by higher levels of cognition, including memory [13]. By using swatches to present colors to study subjects, the WCS researchers attempted to remove any contextual influences on color naming. John Lucy, in [14], though, argues that the three dimensions of color presented in the WCS stimulus array are insufficient for color naming. Namely, they were lacking in degrees of luminosity, reflectance and luster. Furthermore, he argues that color-naming can never be fully detached from referential context and range.

Most natural kinds which people classify have distinct borders of membership. Not so with color. Children only start using color terms with their full referential meaning between ages four and seven despite being able to discriminate colors in dimensions of hue, saturation and brightness [4]. The categorization we take for granted is a hard problem.

2.2.2 Computational Models

There have been a few computational models for color naming. Mojsilovic in [29] describes a system to name an image's dominant colors. The image is first segmented into regions by color and texture, then each color region is named by taking the region's CIE $L^*a^*b^*$ color value and finding the closest prototype in the ISCC-NBS[3] dictionary [24] using a distance metric based on Euclidean distance. Nurminen et al. [32] also name the dominant colors of an image. Image pixel values are converted to CIE $L^*a^*b^*$ space, then

[3]The National Bureau of Standards Inter-Society Council

clustered by k-means and agglomerative clustering. Names are assigned to cluster centers by using unmodified Euclidean distance metric to find the nearest color prototype in a dictionary. An open-source javascript based color naming tool by Chirag Mehta [26] uses a dictionary of color terms combined from wikipedia, Crayola and others.[4] The distance metric used combines RGB values as well as hue, saturation and lightness.

Lammens [22] uses a Gaussian model to select the best color term in a neurophysiologically-derived color space (NPP). He describes a way of combining color labels near the border between color categories to make complex color terms such as "reddish-yellow" and "somewhat blue".

Steele and Belpaeme's target article [39] about getting artificial agents to coordinate color categories by communication (see also [2]) included a color-naming model related to Lammens. The agents simulated in this experiment categorized colors in CIE $L^*a^*b^*$ by using adaptive networks of locally reactive units, a system similar to radial basis function networks. Units of a network have a peak response at one specific color, with exponential decay around it; the final output of a network is the sum of the individual units. Each color category is represented by a network and a categorization made by the network whose response is highest.

Recently, Mengaz et al. [28] demonstrated a model in which each point in the gamut of the OSA uniform color samples is assigned fuzzy membership to the eleven basic color terms. Membership values were assigned experimentally for the OSA samples and interpolated for other points in the space.

One of the problems for all of the above computational color-naming models is that none take into account human color-constancy. It can be argued that white balancing can implemented as a preprocessing step before submitting a color to be categorized, but the color representations chosen for each of the above models attempt to lock colors to specific physiological responses, so preprocessing the image in a sense betrays the impetus for each respective color representation. An alternative representation is the CIE

[4] A list of different color name dictionaries can be found at `http://www-swiss.ai.mit.edu/~jaffer/Color/Dictionaries.html`

CAM color appearance model [30], which attempts to model the perceptual effects of surround, adaptation, illumination and white-point, predicting the *appearance* of a given color. Even with perceptual effects accounted for by white-balancing or a color appearance model, none of the above color-naming models take into account the semantic context of the color being named, something this thesis hopes to address.

2.3 Concept Spaces, Context-Sensitivity and Linguistic Hedges

Peter Gärdenfors proposes a three-layered model of cognition in [11] split between Associationist (connectionist), Conceptual (geometric) and Symbolic (propositional) representations. The central, geometric, component Gärdenfors names conceptual spaces. Abstract concepts, such as *robin*, can be represented as a high-dimensional region in a geometric space with dimensions such as "can-fly" and "has-wings". The region representing *robin* lies within the region for *bird*. Reasoning and inference about concepts can then be transformed into a geometric problem where geometric algorithms can be applied [12]. Conceptual spaces have been applied to both text [38] and vision [7] problems.

In Gärdenfors' model, context effects can be seen as a selective scaling of the conceptual dimensions. On the farm, the concept for bird would scale up the visual "has-wings" dimension, while at the dinner-table, the "tasty" dimension would be emphasized. Applied to color-naming, the context of *wine* would scale the salient color dimensions to bring a deep purple into the region labeled "red".

To communicate about concepts in the word, we must have a shared common ground with our conversational partner. Sometimes, though, it is difficult to determine this shared conceptual space, especially if either the two partners' models greatly differ, or if the word used refers to intangible or invisible things. Arriving at a shared conceptual understanding is the subject of linguistic and cognitive research [6, 15, 9]. Related to this work are linguistic hedges [21], using fuzzy terms like "somewhat brown" or "reddish" to attenuate the meaning of a word or phrase. Hedges are frequently used in referential negotiations. The Steele and Belpaeme target article mentioned earlier connects many of the concepts discussed here: colors are classified by independent artificial agents, who come to a shared understanding of color terms through communication.

2.4 Meta-Classification Techniques

There is a fair body of research about techniques for combining classifiers to increase their predictive power. This class of techniques, in which base classifiers (sometimes called classifier stubs or weak learners) is called meta-classification. The most straightforward of these techniques is voting [20], wherein a number of stub classifiers each make an independent classification and the majority class is chosen as a final result.

Stacking [42] is a generalization of voting where each stub classifier is assigned a weight, and final classification is a result of the weighted vote of the stubs. The weights assigned to the stubs are chosen to minimize error in cross-validation. Stacking is a batch-learning technique due to the weight selection by cross-validation. Bootstrap aggregation (Bagging) [5] creates multiple copies of the training set by drawing samples with replacement. These new training sets are used to create a cohort of stub classifiers whose majority vote is reported as the final classification. Bagging is essentially a smoothing technique, averaging stub classifiers whose decision boundaries are sensitive to training data. Another technique which replicates data is the *Decorate* algorithm [27]. In this approach, data with fuzzy class labels is artificially generated from the training set. This artificial data is used to train stub-classifiers which are combined by voting.

In boosting [36, 37]), each iteration of the algorithm adds a weak learner trained on a weighted dataset, where those examples misclassified by the previous iteration are more strongly weighted. There are a variety of algorithmic variants of boosting, best known of which is perhaps AdaBoost [10].

All of the techniques mentioned above are batch learners. A labeled training set is processed to create a meta-classifier, which remains static for all future classifications. To process new training data, these classifiers must retain their entire original training set. The algorithm described in this thesis does not suffer from this drawback—learning occurs incrementally rather than in batch.

CHAPTER 3

THE CONTEXT DREAMING ALGORITHM

This chapter describes the Context Dreaming algorithm in detail, discussing its operation, critiquing its model, describing its theoretical performance and discussing variants of the algorithm.

Context Dreaming is designed to take advantage of a particular interaction model: one of discrete "interrogations." An example will help clarify what I mean. Imagine an automatic telephone troubleshooter for a computer company. A customer calls and describes a problem with a recently purchased product. The automatic troubleshooter can be seen as a sophisticated classifier, asking questions of the customer and listening to the complaints in order to find the most accurate classification of the problem. Ideally, we'd want the automatic troubleshooter to learn from customers, both within the bounds of a single call (by cup-holder, the customer means compact disc tray) and by aggregating many calls (a whirring noise and smell of burnt hair is likely a power-supply problem). Essentially, there is local, in-dialog fast adaptation where joint definitions are negotiated ("The cup-holder." "The CD tray?" "The thing that slides in and out." "Okay.") and global learning, where the results of multiple conversations are aggregated to help speed the diagnostic process and obtain more accurate results in future dialogs.

This interrogative interaction model is common in real life, and in fact is critical whenever two parties are referring to a shared concept or item. What you mean by "democracy" is likely subtly different from what I mean by "democracy." If you use the term in a way I find surprising, I can ask you to clarify and update my local definition for our conversation. My personal interpretation can remain intact, but we can continue with a shared common understanding. The next time "democracy" comes up in conversation between us, I can recall our shared meaning and proceed without confusion.

A more concrete example—one which motivates the evaluation described later—is that of two parties negotiating the meaning of color terms. Imagine you are sitting across a table from anther person. On the table are two objects whose colors you would describe as cyan and purple. Your interrogative partner says, "Hand me the blue one." Which one did he mean? For you, there is no clear example of a "blue one" so you are forced to decide between the two objects present. Let's say you hand him the cyan object and get the reply, "Thanks." You've now learned that for purposes of this interrogation (and perhaps for future conversations with this partner) colors that you classify under the term "cyan" can also be classified as "blue." An understanding of color has been negotiated.

The dialog model for classification is intimately tied to the functioning of the Context Dreaming algorithm. There are two distinct phases of operation, one which occurs before and during a dialog, and one which occurs afterward. The first, the online *wake* cycle, is analogous to the automated troubleshooter's conversation with a single customer or the negotiation of blue and cyan colors. This phase has a beginning and end, and its duration is much shorter than the lifetime of the classifier (which can continue indefinitely). During the second phase, the offline *sleep* cycle, knowledge learned during a dialog is incorporated into the global model by replaying any new training examples. For our earlier example, it's here that the troubleshooter will generalize from "this customer calls the CD tray a cup-holder" to "sometimes, customers will call the CD tray a cup-holder."

3.1 Context Dreaming as Meta-Classifier

There are a number of techniques that can be used to combine machine classifiers in ways which improve performance, both in speed and accuracy. Perhaps the simplest example is a voting classifier. In this meta-classifier, a collection of sub-classifiers (either heterogeneous or homogeneous), examines an incoming feature vector and performs a classification. The sub-classifiers are sometimes called classifier stubs or stub-classifiers. For a given feature set $\vec{x}$, all the result classifications reported by the stubs are combined by vote; typically, the majority class is considered the winner and final classification. If the component classifiers also produce confidence values with their classifications, then the voting can be weighted accordingly, with more confident classifiers having their votes count more toward the final result. Likewise the contribution of each classifier to the final result can be weighted by another heuristic.

The voting classifier is an example of a *meta-classifier*. Context Dreaming is such a classifier. By consequence, the performance of Context Dreaming is bound to the performance of the base learners within it. A better performing stub classifier will result in a better performing Context Dreaming meta-classifier.

3.2 The 50,000 Foot View

A Context Dreaming classifier contains a library of context-classifier pairs, where the contexts represent the "background information" for a dialog, and the classifiers are any online learner (i.e. a stub classifier). These contexts can be spare—capturing just a subset of relevant contextual clues—rich, or empty. The classifier paired to each context in the library is trained for circumstances appropriate for that context.

At the beginning of a dialog, the Context Dreaming classifier makes a guess as to which classifier would be most appropriate for the given situation, by finding the best context-match in the library for the situation's context. That best-guess classifier is then used for classifications in the dialog. Fast adaptation occurs by heavily weighting new training examples.

If new training examples are offered during a dialog, that dialog's classifier will be reshaped. How can this new knowledge be integrated back into the master Context Dreaming classifier at the end of the dialog? The dialog's reshaped classifier is compared against all stub-classifiers in the library, and the most similar match sequestered. If the highest similarity score is above a threshold, then the training examples gathered during the dialog are *played back* to the sequestered classifier and the situational context merged with the sequestered-classifier's context. If, on the other hand, the score is below the threshold, then the situation's context and the newly reshaped classifier are added to the library.

3.3 Terminology, Parameters and Structure

A Context Dreaming classifier begins with a context c. It then takes a feature vector $\vec{x}$ and classifies it into one of n classes $C_1 \ldots C_n$.

There are six parameters to the Context Dreaming algorithm: two numbers defining a threshold and weight, and four functions for classification, comparison and context merging. These parameters are summarized in Table 3.1. How these parameters are used is explained in the sections below.

The stub classifier which Context Dreaming uses is the first parameter to the algorithm. This classifier must be an online (incremental) learner, and must support weighted learning, where some examples are more important than others. A simple way to add this weighting parameter to a classifier which doesn't have it is to repeat training examples multiple times. In this document, F will represent the class of classifier used as a stub, and f will represent an instance of this stub.

Context Dreaming requires two comparator functions, one for comparing contexts and one for comparing classifiers. Both comparator functions return a similarity score ranging between zero and one, with zero being completely dissimilar, and one being a perfect match.

Another function required by Context Dreaming is M, which merges two contexts into a third.

Label	Function	Constraint
F, f	The stub classifier.	$f(\vec{x}) \rightarrow C \in \{C_1, \ldots, C_n\}$ f is an online learner
S_f	A similarity metric comparing two stub classifiers.	$S_f(f, f') \rightarrow [0, 1]$
S_c	A similarity metric comparing two contexts.	$S_c(c_i, c_j) \rightarrow [0, 1]$
M	Context merging function.	$M(c_i, c_j) \rightarrow c$
γ	A threshold for classifier similarity.	$0 \leq \gamma \leq 1$
ω	A reweighting parameter.	$0 < \omega$

Table 3.1: Parameters to the Context Dreaming algorithm and their constraints.

Finally, two numeric parameters complete a Context Dreaming classifier. A number between zero and one serves as the threshold for classifier similarity (γ). A weighting parameter, ω, sets the adaptation rate during the wake phase.

Structure

A Context Dreaming classifier is a tuple $\langle S_f, S_c, M, \gamma, \omega, L \rangle$ where L is a set of context and stub-classifier pairs, initialized to be empty. During operation, the library is filled with context-classifier pairs, each context in the pair encapsulating the relevant components of the context which best match the paired classifier.

The data structure which describes situational contexts can come in many forms. The version implemented for this thesis is a key-value mapping, where the key is some symbolic label, and the value is a set of strings. In the phone-based troubleshooter example described above, one might choose the context keys such as "caller area code", "time of day", "weeks since product release" etc. For the color-naming experiment described in the next chapter, the context included the participant's native language, age, unique id, displayed image or word, etc. The relevance of any particular key is discovered by the algorithm.

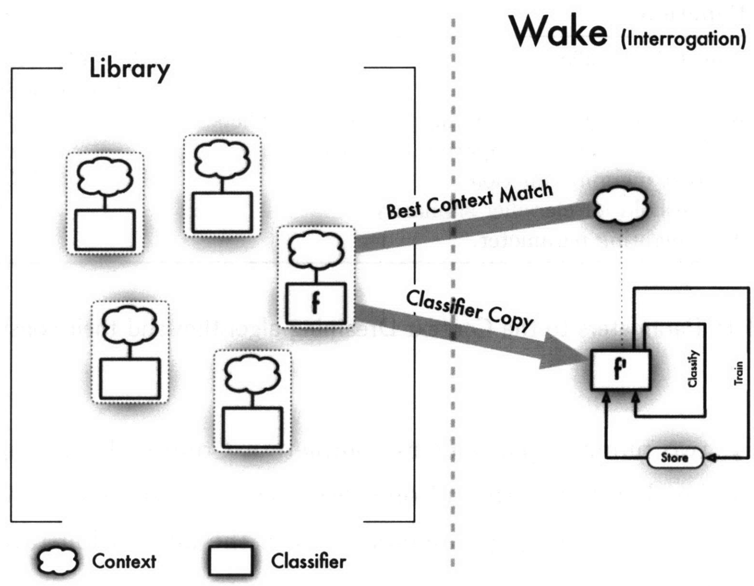


Figure 3.1: Boxology of the Context Dreaming Wake cycle algorithm.

3.4 Phase I – The Wake Cycle

Each wake cycle covers an interrogation with constant context. The beginning of the interrogation is marked by submitting a context data structure to the classifier. This sets the internal state of Context Dreaming for the duration of the interrogation. After submitting the context, any number of classification or training requests can be made as long as the context remains fixed. At the close of the interrogation, a signal is sent to the Context Dreaming classifier, ending the wake cycle. A single wake cycle corresponds to a single interrogation.

The submission of a context (c) to Context Dreaming primes the classifier. First, Context Dreaming iterates over all the context-classifier pairs in its library L, comparing them to the incoming context using the context comparator S_c. The context receiving the highest score when compared to c is selected along with its accompanying classifier. Call this pair $\langle c_{max}, f_{max} \rangle$.

If the library is empty, then c is used as c_{max} and the classifier prototype F is used as f_{max}.

Next, a copy of this maximum scoring classifier is made (f'_{max}) and set aside. The library L remains intact during the wake cycle. All classification and training examples submitted to Context Dreaming for the duration of the wake cycle are passed through f'_{max}. Training examples are submitted to f'_{max} with the weighting parameter ω. They are also stored for replay during the sleep cycle. It is by this means that a custom classifier is trained for the duration of the interrogation. In analogy to the hypothetical example, I learn what you mean by "democracy".

Algorithm 1 The Context Dreaming Wake Cycle Algorithm

On input $\langle c \rangle$:
if (L is empty) **then**
  $\langle c_{max}, f_{max} \rangle = \langle c, F \rangle$
else
  $\langle c_{max}, f_{max} \rangle = \underset{\langle c_i, f_i \rangle \in L}{\text{argmax}} \left(S_c(c, c_i) \right)$
end if
$f'_{max} \leftarrow \text{copy}(f'_{max})$
while (The interrogation is active) **do**
  if (Request is for a classification) **then**
    Return the result of $f'_{max}(\vec{x})$
  else if (Request is a training example $\langle C_i, \vec{x} \rangle$) **then**
    Train f'_{max} with $(\omega, C_i, \vec{x})$
    Store example $\langle C_i, \vec{x} \rangle$
  end if
end while

3.5 Phase II – The Sleep Cycle

At the end of an interrogative wake cycle, the Context Dreaming algorithm incorporates what it learned for future use. During this phase, the stub classifier f'_{max} that was retrained over the course of the interrogation is incorporated into the library. The integration happens in two steps. First, Context Dreaming uses the classifier comparator S_f to compare f'_{max} (the retrained classifier used during the wake cycle) against all classifiers currently in the library, L.

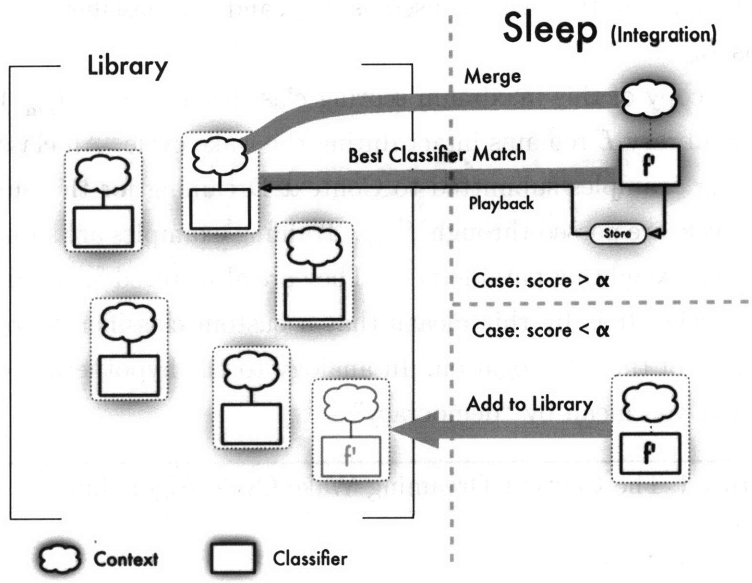


Figure 3.2: Boxology of the Context Dreaming Sleep cycle algorithm.

Once the closest match is found, Context Dreaming completes the integration. Consider the library classifier and associated context with the highest classifier similarity score s: $\langle f_i, c_i \rangle_{max}$. If the score $s > \gamma$ then the training examples gathered during the wake cycle are "replayed" for f_i, training f_i using a weight of one. The two contexts c_i (paired with the library classifier) and $c_{f'}$ (paired with the interrogation's context) are merged together using M. This merged context is used as the new key for f_i.

Otherwise, if $s \leq \gamma$ then the f'_{max}, and its associated context $c_{f'}$ are added to the library.

3.6 Expected Performance and the Effect of Parameters

Making claims about the theoretical performance of a Context Dreaming classifier is difficult because of the wide flexibility of choosing a stub classi-

Algorithm 2 The Context Dreaming Sleep Cycle Algorithm

On input $\langle c_{f'}, f'_{max}\rangle$ {Wake-cycle classifier f'_{max} and the interrogation context $c_{f'}$}:
$\langle f_{match}, c_{match}\rangle \leftarrow \underset{\langle f_i, c_i\rangle \in L}{\text{argmax}} \left(S_f(f_i, f'_{max})\right)$
$bestscore \leftarrow \max\left(0, S_f(f_{match}, f'_{max})\right)$
if $bestscore > \gamma$ **then**
  $L.remove\left(\langle f_{match}, c_{match}\rangle\right)$
  for Training example $\langle C, \vec{x}\rangle$ **do**
    Train f_{match} with $\langle C, \vec{x}\rangle$ and weight 1
  end for
  $c_{merge} \leftarrow M(c_{f'}, c_{match})$
  $L.add\left(\langle c_{merge}, f_{match}\rangle\right)$
else
  $L.add\left(\langle c_{f'}, f'_{max}\rangle\right)$
end if

fier, context data type, comparators, and the numeric parameters. Nevertheless, some trends based on the effects of the parameters can be expected. As with other meta-classifiers, the performance of Context Dreaming is dependent on the performance of the stub classifier. We can expect that Context Dreaming will perform as well as the stub, but this is not guaranteed. In fact, if the γ is set low, then no new stub classifier will be added to the library—all training examples will be shunted to the prototype stub classifier F. Essentially, when γ is very small, then Context Dreaming reduces to the stub classifier but with in-dialog fast adaptation. Over-fitting will result if γ is set too high. In that case, the library will fill with contexts and classifiers that will be infrequently used.

The time-performance of Context Dreaming can be predicted as a function of the performances of the parameter functions. The startup time of the wake cycle is $O(|L| \times O(S_c))$ because of the single loop through each of the library's contexts. Any classifications and training during the wake cycle are $O(f_{\text{classify}})$ and $O(f_{\text{train}})$ respectively: Context Dreaming merely passes the feature vector to the selected stub-classifier or adds a constant-time storage of training examples. Sleep-cycle time performance is not much different: $O\left(|L| \times O(S_f) + m \times O\left(f_{\text{train}}(\vec{x})\right)\right)$ with m being the number of training samples collected during the wake cycle. During this offline part, there is a single loop through the library, comparing classifiers, followed by a training round looping once through the examples.

3.7 Algorithm Variants

The Context Dreaming algorithm provides fodder for a number of variants. Three are discussed here, the first of which may address some concerns about stability, the second scalability, and the third which can make more efficient use of the training data under certain assumptions of the data's form. Many other refinements to the algorithm can be imagined, whether conceptual or in implementation.

3.7.1 Hedging Your Bets

The Context Dreaming algorithm makes a hard guess by selecting a single stub classifier to take part in the wake cycle. If multiple contexts in the library receive the same top score when compared against the situation's context, there's no guarantee that the stub classifier Context Dreaming will choose will be correct. One way to soften this hard guess, and effectively have the meta-classifier hedge its bets is by choosing the top k context-classifier pairs from the library. These top k classifiers would vote to decide on a final classification for a feature vector $\vec{x}$. Voting could be weighted by each classifier's respective context-similarity score and classification confidence (if the stub classifier returns a confidence score.)

The sleep-cycle is also modified for this variant. Training examples gathered during dialog are reclassified by each of the k stubs and used to get a post-hoc evaluation of whether that classifier should have been included in the voting cohort. Those stubs which score above a threshold would be integrated into the library as described above. Those below would be discarded.

This modification should make Context Dreaming more robust and reduce the variance of its classification error rate. The post-hoc assessment decreases the chances that a stub classifier in the library would be trained for a situation inappropriate for its context.

3.7.2 Refined Context Intersection

The Context Dreaming algorithm is agnostic to the description of context as long as the context comparator and intersection function match their respective constraints. The version implemented to demonstrate Context Dreaming operation though is limited by the context-merging and comparison functions—merging is accomplished by returning a context containing the intersection of the input contexts, and scoring is also based on amount of overlap. Therefore, merged context can only represent joint existence in the context ("ands"), with no way to represent alternatives ("ors"). The refined context intersection described here is intended to overcome some of the first iteration's limitations.

The refined context is represented as a key-value histogram. Each value in the context is augmented with a count. Contexts are intersected by summing the counts in the values.

$$\left\{ \begin{array}{rl} \textit{image}: & \langle \textit{grapes}: 1\rangle \\ \textit{language}: & \langle \textit{english}: 1\rangle \end{array} \right\} \quad \text{and} \quad \left\{ \begin{array}{rl} \textit{text}: & \langle \textit{eggplant}: 2\rangle \\ \textit{language}: & \langle \textit{english}: 2, \textit{japanese}: 1\rangle \end{array} \right\}$$

are merged into

$$\left\{ \begin{array}{rl} \textit{image}: & \langle \textit{grapes}: 1\rangle \\ \textit{language}: & \langle \textit{english}: (1+2), \textit{japanese}: 1\rangle \\ \textit{text}: & \langle \textit{eggplant}: 2\rangle \end{array} \right\}$$

Using histograms for context values allows for better context similarity scoring. The context comparator function can produce a fuzzy notion of "and" as well as "or" using a relative entropy score such as the Kullback-Leibler divergence.

3.7.3 Training a Filter

This variant of Context Dreaming allows training examples to be applied to *all* contexts within the library and embraces Gärdenfors' Conceptual Spaces [12]. To accomplish this, the Context Dreaming classifier is modified, adding a parametric feature-transformer $g(\theta, \vec{x}) \rightarrow \vec{y}$, where θ are the parameters of the transform. Furthermore, the library of context and stub-classifier pairs is replaced by a library of context and feature-transformer pairs. The wake and sleep cycles are changed as follows:

In the wake phase, the best matching context is chosen as described above. Any request for classification is first passed through the chosen feature-transformer, then classified by the stub classifier F. Fast adaptation for the duration of the wake phase comes by learning the parameter θ (by hill climbing, simulated annealing or other such technique).

At the end of the interrogation, the newly trained feature transform is integrated into the library as is described above. Rather than a classifier comparator S_c, this variant uses a transform comparator $S_\theta(\theta_i, \theta_j) \rightarrow [0, 1]$ to score transform similarities. The γ parameter now applies to this similarity score. Any training examples gathered during the wake cycle are played back though $g(\theta, \cdots)$ and used to train the single stub classifier F.

3.8 Discussion

Comparing Context Dreaming to other machine learning algorithms can yield the following critique: How is Context Dreaming different from other mixed-data-type classifiers? Can't the contextual information be incorporated into a single feature vector? Essentially:

$$
\begin{aligned}
&\vec{x} = \{x_1, \ldots, x_n\} && \text{where} \\
&\vec{x}_{context} = \{x_1, \ldots, x_i\} && \text{and} \\
&\vec{x}_{features} = \{x_{i+1}, \cdots, x_n\}
\end{aligned}
$$

My response is to focus on the particulars of the use of a Context Dreaming classifier. Essentially, Context Dreaming should be considered *within the context of its use.* Context Dreaming takes advantage of having a static component (the context) and a dynamic one ($\vec{x}$). The algorithm "locks in" on a particular stub classifier for the duration of an interaction: this fact allows for local adaptation *to a particular interlocutor* in a way that is not possible with a more general classifier. Furthermore, as a system, Context Dreaming is straightforward and flexible. It allows classifiers that use only one data type (e.g. numeric values) to be augmented with mixed data types (e.g. symbolic contexts).

I conclude this chapter by summarizing the advantages of Context Dreaming and the ways it takes advantage of the dialog model it works in.

- Context Dreaming allows for fast adaptation during a dialog with fixed context.
- The two phases of operation allow Context Dreaming to provide fast answers during an online dialog, and shunt more computationally expensive procedures to the offline sleep cycle.
- Context Dreaming is well suited to interrogative tasks—situations which frequently arise in dialogues where there is a negotiation of the meanings of words.
- Classifiers accepting a single data type are transformed into mixed data type classifiers.

The next two chapters describe a color naming experiment and an evaluation of Context Dreaming on the corpus gathered.

CHAPTER 4
CONTEXT EFFECTS ON COLOR NAMING

The words we use to label colors in the world are fluid. They are dependent on lighting conditions, on the item being named, and on our surroundings. The color stimulus you might label as "orange" in one context, you would label "red" when talking about hair. Likewise, "black" becomes "red" when talking about wine. The grass would still be called green when lit by a red-tinted sunset. Although we intuitively know this context effect exists, I wish to quantify it under controlled circumstances.

This chapter describes an experiment I designed in part to gather a corpus on which to evaluate the Context Dreaming algorithm. The experiment was built upon a particular color negotiation task described previously. A colleague sitting across a table asks you to "pass the blue one." To your eyes though, there's only a cyan object and a purple one. Which do you choose? The experiment described here distills this task to its most primitive components. Further discussion of the way the experiment encodes this hypothetical scenario can be found in the next chapter, which describes the application of the corpus on a Context Dreaming classifier.

The results of the experiment confirm that semantic context affects color categorization, although sometimes in surprising ways. The first part of this chapter describes the experiment performed, and the second discusses the

sex	#	Min age	Max age	Mean age
Male	8	18	45	26.6
Female	15	18	63	43.3

Native language	#
English	18
Chinese	3
Portuguese	1
Spanish	1

Table 4.1: Demographics of the study participants.

results and proposes a model which may account for the data.

4.1 Experiment

I designed an experiment to validate the hypothesis that situational and semantic context affect the naming of colors. The experiment consists of three color-related tasks: calibration, forced choice and naming. The calibration task provides a baseline on which to evaluate the naming and forced choice tasks. Both naming and forced choice parts evaluate contextual effects on color categorization by presenting an ambiguous color stimulus and forcing the experimental subject to make a categorical decision.

To prepare stimuli to be presented in this experiment, a separate stimulus-selection data collection was run.

4.1.1 Participants

Thirty-six participants were solicited from the MIT community by email announcements and posters. Inclusion criteria was proficiency with the English language. Participants were asked to provide their age, sex, native language and any other languages they spoke fluently. Participants were compensated for their time. From these, the first 13 were chosen to complete only the stimulus-selection task.

4.1.2 Equipment

The experiment was performed in a windowless, dimly lit room illuminated at approximately $3200°K$. Approximately ten minutes were spent adjusting

Chromatic	Achromatic
red	black
green	white
blue	gray
yellow	
orange	
purple	
brown	
pink	

Table 4.2: The eleven basic color terms of the English language, as recorded by the World Color Survey[19]. The chromatic terms were used in the color context experiment described here.

to the ambient lighting in the room before any color-related tasks were performed. Color stimulus was presented with custom-written software on an Apple Macintosh computer and data recorded to a relational database. A 30-inch Apple Cinema Display was used as the display device. The monitor was calibrated to the sRGB standard (D65, 2.2 Gamma) using a ColorVision Spyder2Pro hardware color calibrator.

All color stimulus was presented against a neutral gray background.

4.1.3 Stimulus-selection

All of the experiment tasks described below share a common set of stimuli colors chosen as follows: A pair of colors are selected from the eight basic chromatic color terms of the English language (see Table 4.2). A third color is chosen "in between" the first two. The idea is to make this mid-point color as categorically ambiguous as possible so that a participant, having to make a choice to fit the color to a basic color term would have the most difficult time. Essentially, these mid-point colors lie on the decision boundary between two color terms.

The ambiguous colors were chosen experimentally. Thirteen participants were presented with pairs of color terms and asked to find the most ambiguous mid-point color. Stimulus was presented under the same experimental conditions as for the full experiment.

The set of color terms C contains the focal colors for the eleven basic color terms found in [19]. This set of CIE L*a*b* triples were taken from the World Color Survey data archive.[1] For each color pair ($\langle C_i, C_j \rangle$), a rectangular swatch was presented flanked by color terms. On the left, the term for color C_i; on the right, color C_j's term. The ambiguous color X filled the center swatch. Below the swatch and color labels, a slider allowed the participant to change the mix between the two colors. The CIE L*a*b* value of the ambiguous center color was determined as:

$$L^*_X = \frac{L^*_{C_i} + L^*_{C_j}}{2}$$
$$a^*_X = \alpha a^*_{C_i} + (1-\alpha) a^*_{C_j}$$
$$b^*_X = \alpha b^*_{C_i} + (1-\alpha) b^*_{C_j}$$

where α is the slider value, which ranges over $[0, 1]$. By allowing only variation in a and b, only the chromaticity of X varies. I chose to fix the luminance of the ambiguous color in order to minimize biasing based on perceived brightness and minimize perceptual contrast effects due to the experimental stimuli being presented against a neutral gray background.

Values of α for each of the color pairs was gathered from thirteen participants. From this data, the mean ($\mu_{\alpha_{ij}}$) and variance (${\sigma^2}_{\alpha_{ij}}$) were calculated. Trials for the full experiment described below used ambiguous colors derived from the statistics $\mu_{\alpha_{ij}}$ and ${\sigma^2}_{\alpha_{ij}}$. Table 4.3 summarizes the statistics gathered.

4.1.4 Color-Survey Task

The color-survey task consisted of two components similar to the tasks of the World Color Survey: labeling focal colors and color-class membership. In both cases, a palette of colors approximating the World Color Survey stimulus was presented. For each of the eleven basic color terms of English, participants were asked to select both *the most representative color swatch for that label* as well as *all color swatches covered by that label.* CIE L*a*b*

[1] http://www.icsi.berkeley.edu/wcs/data.html

C_i	C_j	α	σ_α	C_i	C_j	α	σ_α
red	green	0.470307	0.043355	blue	orange	0.416461	0.051146
red	blue	0.534923	0.074149	blue	pink	0.617692	0.099827
red	yellow	0.419846	0.200738	blue	brown	0.545307	0.031639
red	orange	0.396538	0.200362	blue	purple	0.558076	0.047334
red	pink	0.267307	0.157634	yellow	orange	0.208461	0.190263
red	brown	0.434461	0.102883	yellow	pink	0.559230	0.091499
red	purple	0.371615	0.098075	yellow	brown	0.624846	0.384815
green	blue	0.509307	0.051746	yellow	purple	0.602615	0.057751
green	yellow	0.512692	0.079417	orange	pink	0.527846	0.094767
green	orange	0.475153	0.068584	orange	brown	0.565461	0.119777
green	pink	0.554769	0.053107	orange	purple	0.519076	0.059021
green	brown	0.746384	0.072170	pink	brown	0.540000	0.076685
green	purple	0.580153	0.030303	pink	purple	0.326769	0.133534
blue	yellow	0.359923	0.047062	brown	purple	0.465307	0.064446

Table 4.3: Results of the ambiguous-color calibration task.

values for each of the swatches were taken from the World Color Survey data archive[2]. Unfortunately, many of the colors of the WCS stimulus lie outside the gamut of sRGB. Those colors were converted from CIE L*a*b* to sRGB, then clipped at the maximum RGB value.[3] As a result, the stimulus presented in this experiment is not a complete analogue to the WCS and therefore direct comparison to the WCS data is problematic. Nevertheless, the clipped values presented for this experiment are sufficiently spread across sRGB space to provide a measurement of contextual effects on naming.

The choice to use the WCS stimulus set was made in part because of the singular prominence of the WCS in color-naming research. Despite the disparities between the sRGB stimulus presented in this experiment and the swatches used in the WCS, comparing the data gathered against the WCS data provides a certain degree of confidence that the participants in the study are sufficiently "typical."

[2]http://www.icsi.berkeley.edu/wcs/data/cnum-maps/cnum-vhcm-lab-new.txt
[3]For the Context Dreaming evaluation, this clipped value was used.

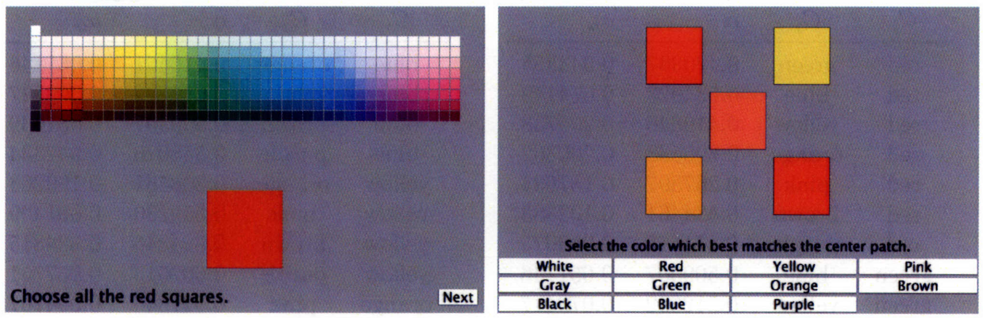


Figure 4.1: On the left, an example of a color-survey task. To the right, an example of the naming task. The colors appearing here will vary from the stimulus due to color variation in the printing or display of this figure.

4.1.5 Binary Forced Choice Task

In forced choice tasks, the participant is presented with a color stimulus in the center of the screen and two color labels in black text on the left and right. In order to proceed to the next screen, the participant must choose which of the two labels better represents the center stimulus. Participants were instructed to proceed as quickly as they believed they could make an effective decision. Response time was recorded.

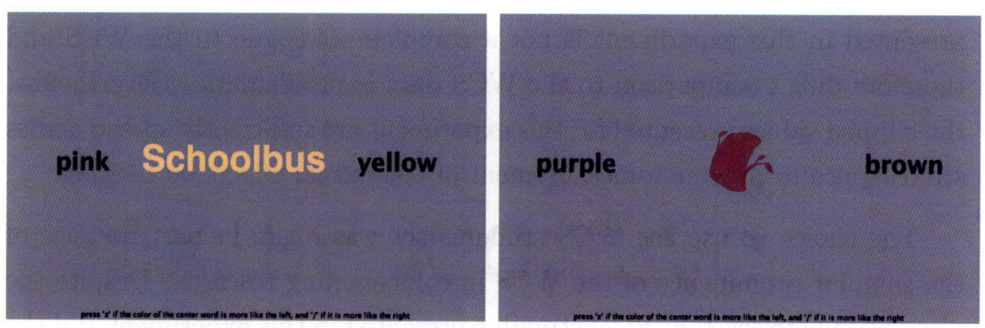


Figure 4.2: Examples of forced-choice tasks; the left image with a word context, the right with an image context. Colors in this document vary from the stimulus due to color variation in printing or display.

In order to help negate any perceptual saturation effects, each stimulus was preceded with one second of the screen at neutral gray. Furthermore, 1500 milliseconds after becoming first visible, the stimulus disappeared, leaving only the color labels.

The three types of binary choice tasks are:

Context-free (Control): The ambiguous color stimulus is presented as a 200×200 pixel color swatch in the center of the screen.

Word context: The word context for a color pair was an extra-bold 96 point font. The color of the word was the ambiguous color X. This is a variant of the Stroop task[40].

Image context: A high-contrast iconic image was used as a stencil when presenting the ambiguous color. A different iconic image was used for each of the eight color terms.

The forced choice tasks were performed on the color pairs shown in Table 4.4. Each pair (C_i, C_j) was presented with five different ambiguous center colors, generated using $\alpha_{ij} = \mu_{ij} \pm \beta\sigma_{ij}$ where $\beta \in \{0, 0.75, 1.5\}$. These five colors were furthermore presented with the label for C_i on the right, the label for C_j on the left, as well as the reverse. For the two tasks testing contexts, the number of decisions was furthermore multiplied by two because contexts for C_i and C_j were presented.

This brought the total number of samples for each color pair to ten for the context-free (control) case, and twenty otherwise. The total number of decisions per participant collected during the experiment was 400 (8 color pairs, 10 or 20 decisions per pair, 3 tasks).

4.1.6 Naming Task (Surround Context)

This task was designed to determine any effect that color classifications of surrounding objects may have on a naming task. It is similar in design to the experiment described in [35]. The stimulus presented to the participant was five square color swatches arranged on a three-by-three grid. The contextual swatches were placed at the four corners of the grid. In the center square was the target stimulus.

As in the previous tasks, two colors (A and B) were chosen from the eight chromatic basic English color terms. Between these two colors, five

		Context			
A Color	*B* Color	***A* word**	***B* word**	***A* image**	***B* image**
Red	Green	Cherry	Broccoli	Cherry	Leaf
Green	Blue	Broccoli	Ocean	Leaf	Waves
Blue	Red	Ocean	Cherry	Waves	Cherry
Pink	Yellow	Flamingo	Schoolbus	Flamingo	Corn
Yellow	Orange	Schoolbus	Carrot	Corn	Traffic Cone
Orange	Purple	Carrot	Eggplant	Traffic Cone	Grapes
Purple	Brown	Eggplant	Chocolate	Grapes	Log
Brown	Pink	Chocolate	Flamingo	Log	Flamingo

Table 4.4: Color pairs used in the experiment and the word and image contexts used. Prints of the images can be found in the appendix.

ambiguous colors were generated as above. One of the two colors (Lets say A) was chosen to provide the context. The four corner colors of the stimulus were filled with variants of A by rotating the hue of A by a fixed amount in either direction. Two cases were tested; one with a maximum deviation of 135° of hue, and one with a maximum deviation of 18° of hue.

At the bottom of the screen were buttons that the participant would press in order to make a classification into one of the eleven basic color terms. Response times were also measured for this task.

4.2 Results and Discussion

The results of the experiment largely confirmed the hypothesis that context affects color categorization. There were some surprises, especially in the binary forced choice tasks.

Color Foci and Classes:

As a measure of inter-rater agreement, I calculated average information entropy for each color. For each sample swatch ($\Box_i$), I calculated $p_i(\Box_i \in C)$ and $p_i(\Box_i = F_C)$, the probability that an annotator labeled it a member of color class C, and the probability it would be labeled the focal color for class

	Mean color class size	Class membership entropy	Focal color entropy
black	1.48	0.00581	0.00000
red	6.04	0.03882	0.00881
gray	6.43	0.01990	0.01029
brown	8.43	0.05024	0.01054
white	10.26	0.08833	0.00156
orange	12.78	0.06934	0.01247
yellow	12.83	0.09403	0.00935
pink	16.74	0.11437	0.01408
purple	17.39	0.10167	0.01341
green	29.26	0.15477	0.01240
blue	50.91	0.17822	0.01368

Table 4.5: Mean color class size and average information entropy for color swatches in the color foci and color class tasks.

C. From this probability, I calculated the mean information entropy (H):

$$p_\in = p_i(\Box_i \in C) = \frac{n_{i\in C}}{N}, \quad H = \frac{1}{m}\sum_{i=1}^{m}\left[-p_\in \log_2(p_\in) - (1-p_\in)\log_2(1-p_\in)\right]$$

where N is the total number of participants. Table 4.2 collates the results. The low entropy indicates strong agreement among the annotators.

Binary Forced-Choice Tasks:

Results from all three binary forced choice tasks are presented in Figure 4.3. Each grouping of five bars represents results for one color pair. For a given color pair A-B, the fraction of times a participant chose the first of the two colors (A) was tabulated and this fraction was averaged across all participants who were qualified. Participants were disqualified if the control case was unambiguous for a given color pair (i.e. the participant consistently chose one of the two colors in the control case).

Consider the leftmost color pair, brown-purple, which had a strong context effect. The bar for the control case shows the ambiguity of the stimulus color. The next two bars show that participants deviated toward choosing the brown category when either the word context was "chocolate" or the im-

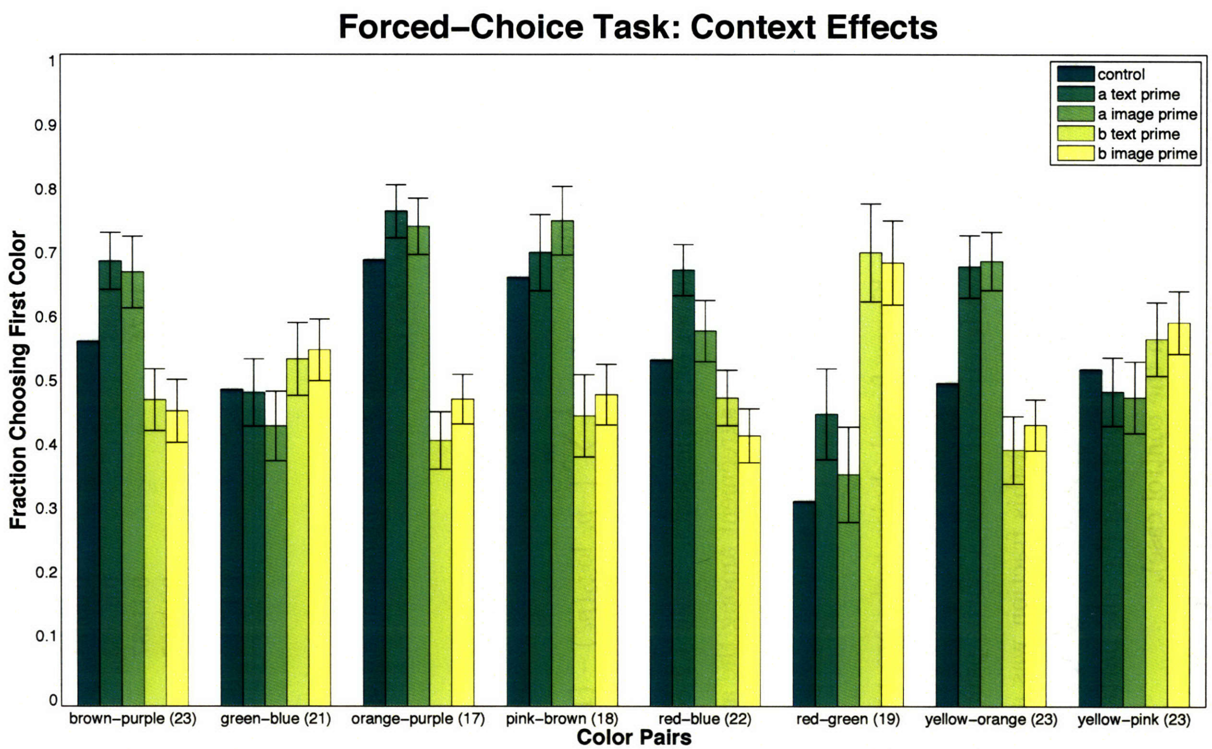


Figure 4.3: Results of the binary forced-choice tasks, grouped by color-pair. The number in parentheses next to each color pair is the number of participants included.

age was of a wooden log. They deviated toward purple when the text context was "eggplant" or the image context was of grapes (the final two bars). An effect was considered positive if an A-context caused A to be chosen more frequently, or a B-context increased the likelihood of B being selected. Similarly, an effect was termed negative if an A-context caused B to be more frequently chosen, or a B-context increased A's likelihood. Positive context effects occurred with five of the eight color pairs. Surprisingly, one color pair (red-green) showed a strong negative context effect and two others—green-blue and yellow-pink—showed slight negative effects. Response times (Figure 4.4 for both context-sensitive tasks were almost identical to the context free task, indicating the context did not introduce new task demands or strategies for the participants.

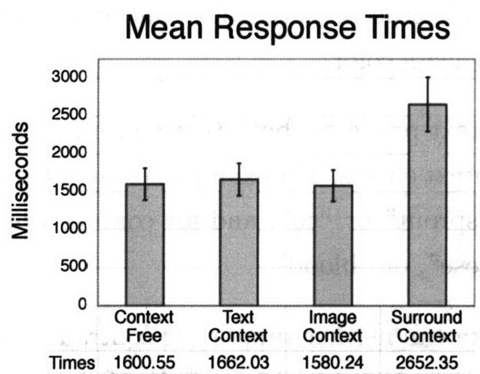


Figure 4.4: Average response times for the binary forced choice tasks and the surround context task. The nearly identical mean response times for the binary-choice tasks indicate that the test was successful in capturing participant's unedited responses. The increase in response time for the surround context task was most likely due to the time required to select a choice using the mouse. For the binary choice, participants made their selection by pressing one of two keys on the keyboard.

I believe that the negative context effects observed with three of the color pairs are caused by color category boundaries varying between context and context-free prototypes, coupled with the specific choices of contexts for this experiment. Consider the schematic in Figure 4.5, which shows hypothetical class boundaries for two colors, A and B. The ambiguous color X is shown approximately half-way between these boundaries along the line connecting A and B's focal colors. The second boundary around the focal color for A demarcates the the colors *prototypical for the image context used for A.*

This boundary is a small subset of the A color task, indicating that for this specific choice of image context, only a small set of colors are deemed typical. Under this set of circumstances, the ambiguous color X is closer to the boundary of B—leading to a negative context effect. The negative effect was observed weakly with color pairs green-blue, yellow-pink and strongly with red-green.

This experiment used only one term for each color as text context and a single iconic image for each color's image context. Two minor adjustments to the experimental procedure would be able to confirm this hypothesis:

1. Adding to the color-survey task requests for the participant to select all colors typical to a given context condition. For example: "Select all the *Flamingo* colored swatches." Armed with this data, I hypothesize that the distance between the ambiguous red-green color would be closer to a red-class color than a broccoli-class color, and closer to a green-class color than a cherry-class color.

2. Adding a larger variety of context primes to both text and image sets should reverse the negative context effects. For example, context words for green such as "grass" or "sprout" or "go", and for context words for red such as "stoplight", "rose", or "blood".

Unfortunately, the current set of data can neither support or contradict the proposed model. A future experiment including either or both of the above procedural modifications is required.

Surround Context Effects:

Figure 4.6 shows results of the surround-context naming task. Data for each color pair A-B was grouped into three categories. A participant's selection for the ambiguous color's name falls into the *congruent*-choice category if it matches the color name of the surround context. Likewise, it falls into the *contrast*-choice category if the name matches the second term of the color pair. If the participant selected a name matching neither color, then it is categorized in the *other* category.

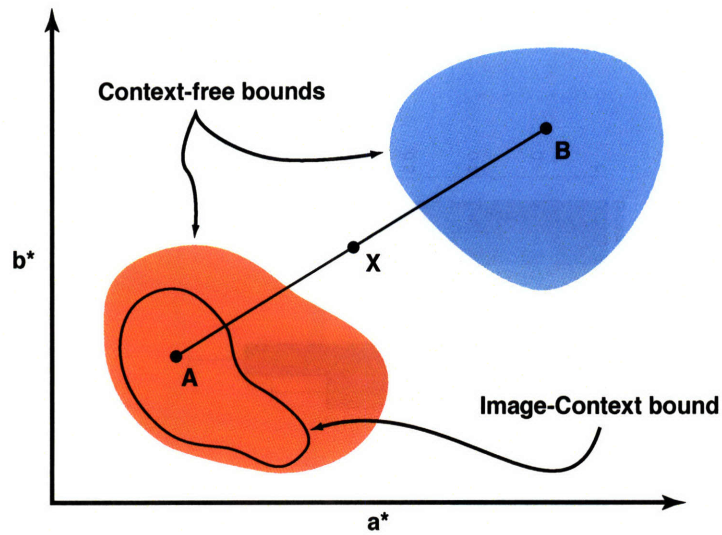


Figure 4.5: Schematic for a model which may explain the binary forced-choice results.

As with the Binary Forced-Choice task, the surround context had a strong effect on color naming, though surprisingly, the effect was sometimes congruent, and sometimes contrasting. The ambiguous color of four of the eight color pairs was most frequently named neither A nor B. Of the remaining four color pairs, two showed a congruent context effect (brown-purple and green-blue) and two showed a contrast context effect (pink-brown and yellow-orange). I believe that these results may be explained by a perceptual contrast effect. The luminosity of the color pairs with congruent effects were equal, compared to the color pairs which showed a contrast effect ($\Delta L^* = 40.81$ for pink-brown, $\Delta L^* = 19.65$ for yellow-orange).

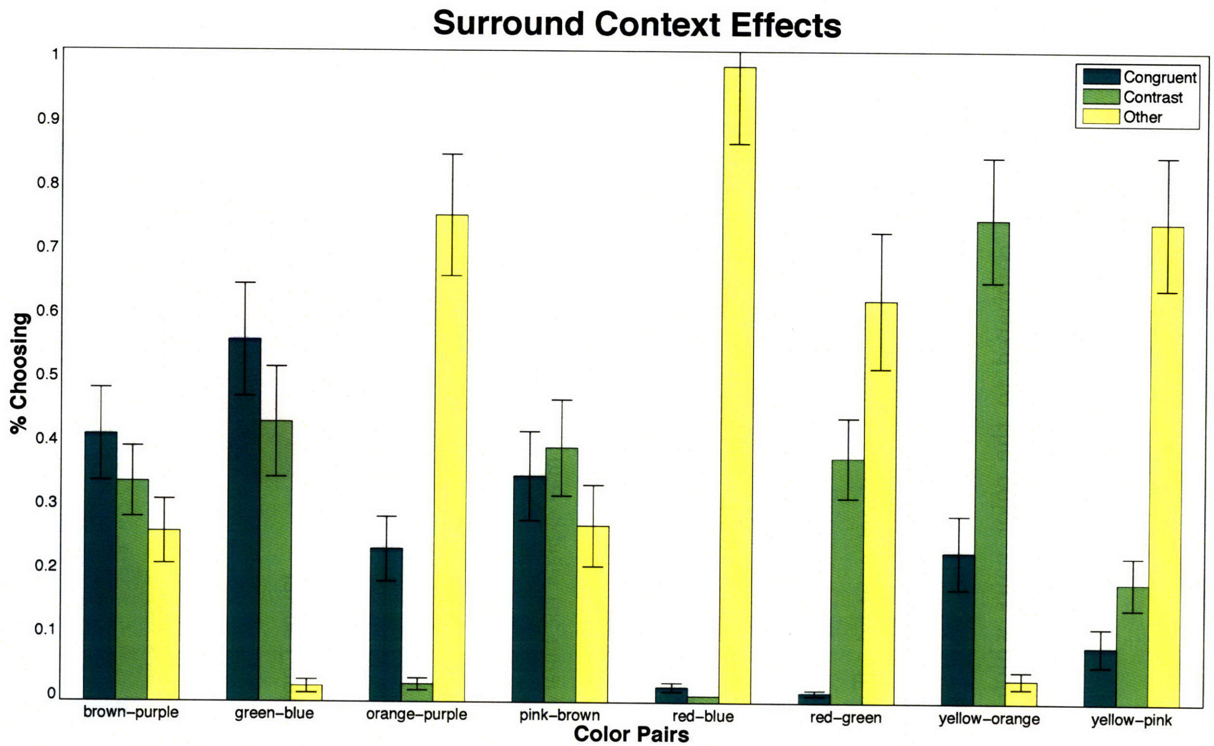


Figure 4.6: Results of the surround-context naming tasks, grouped by color-pair A-B. Results by each color pair are grouped into congruent choice (Context A–Choice A, or Context B–Choice B) or contrast choice (Context A–Choice B, or Context B–Choice A). The third category, "Other" includes all naming choices which were neither A nor B.

CHAPTER 5

EVALUATING CONTEXT DREAMING

How well does Context Dreaming perform? Can it predict human color-naming responses? To evaluate the theoretical performance of Context Dreaming on real world data, I implemented Context Dreaming and used the experimental data described in Chapter 4 to compare it against a baseline. Finding an appropriate evaluation procedure is challenging because to the best of my knowledge, there do not exist other online classification algorithms specialized for the interrogative model Context Dreaming relies on.

To evaluate Context Dreaming, the idea was to find a bare-bones dataset which encapsulates a problem of interrogatory learning while stripping away all components not related to the classification task; speech recognition, natural language processing, computer vision, etc., though necessary components for a holistic system which can participate in dialog, are obscuring factors when trying to evaluate the Context Dreaming classifier.

The problem which first prompted me to explore interrogative classifiers—table-top color negotiation—serves as the evaluation corpus for Context Dreaming. This corpus is a subset of the data from the experiment described in Chapter 4. The binary forced-choice task from the color context experiment in particular provides a dataset which encapsulates table-top color negotiation. How so? First, a brief recap of the task.

Two people are at a table with two colored objects on it. Each person has their own distinct (though similar) model for colors. In conversation, the interlocutors refer to the objects by color label. Problems arise when the color labels one person assigns the objects do not match the second person's labels. To arrive at a shared understanding of color, they must learn new mapping from tristimulus values to color categories—though these mappings will in all likelihood be related to the mappings each party arrived with.

Returning to the evaluation scenario, consider a Context Dreaming classifier which takes the place of one of the interlocutors. At the table are objects X and Y, both of ambiguous color when in a context-free scenario (X between colors A and B, and Y between B and C). A request to "Hand me the B colored one" is made. Using the context of the situation—the category of the objects X and Y, the particular conversation partner, the language being spoken, etc.—which should the classifier select?

The experiment's binary forced-choice task presents a similar problem. An ambiguous color X, at the categorical junction of colors A and B, is presented under a controlled context. The question can then be posed to the classifier: is A or B a better label for X. If the classifier chooses B, and the participant also chooses B, this is analogous to the classifier making the correct decision in the table-top color negotiation task.

The rest of this chapter discusses the implemented Context Dreaming classifier and its evaluation against the color-experiment corpus. The chapter concludes with a discussion of lessons learned.

5.1 The Implemented System

A Context Dreaming classifier was implemented in Java. For this proof-of-concept, all the components of the classifier—the context data type, the stub classifier, and the similarity metrics—were made as simple as possible.

Modified Hedging

The implemented system uses a modified version of the k-best bet-hedging modification described in Section 3.7.1. When entering the wake cycle, the classifier selects the k context-classifier pairs who scored best with S_c. For color triples to be classified, the k classifiers combined their results by weighted voting. A stub-classifier's weight was equal to the context similarity score between its paired library context and the situational context of the wake cycle. When training on new examples, all k stubs were trained. The sleep cycle remained almost identical. One modification was necessary due to the bet-hedging: only the one stub-classifier with top context similarity score was compared against the library classifiers and considered for addition to the library (if its classifier-similarity was less than γ, as per the algorithm).

I chose to use this k-best variation because the choice of context datatype and context-comparator made it likely that the top context-similarity score would be shared by multiple context-classifier pairs. The effect of this modification was to smooth the error rates reported below.

For the evaluation reported here, k was set to five.

Context

Contexts were represented as a sparse map of symbols. Keys in the map were: unique participant id, participant sex, image context, word context, and participant native language. Some context keys which had little or no effect on color decisions, such as participant sex, were intentionally used as confounding factors. The following is an example of a context:

$$\left\{ \begin{array}{rl} id & : 10 \\ sex & : male \\ image & : flamingo \\ language & : english \end{array} \right\}$$

Stub Classifier

A histogram model was used as the stub classifier. Histogram models make no strong claim on the form of the distribution they represent. Mengaz et al. showed that some color categories in CIE L*a*b* are concave [28]. Histograms can effectively capture that observation. Eight bins were used across each dimension for a total of 512 bins.[1] Histograms were smoothed with a radial Epanechnikov kernel spanning three bin widths.

A histogram bin typically contains the count of examples which fall into it, though in order to incorporate weighted training (the ω parameter) this had to be modified. For this refashioned histogram, when an example is added, a percentage of the total "count"[2] in the histogram is added to the bin. Explicitly: say the total "count" in the histogram is n and the "count" in bin i is m. After adding an example to bin i with weight ω, the new "count" in that bin is $m + \max(1, \frac{\omega n}{100})$.

Colors were represented as CIE L*a*b* triples; and the distance between pairs was calculated as the Euclidean distance in CIE L*a*b* space. Other color distance metrics have been proposed to linearize disparities in CIE L*a*b*, notably Mojsilovic [29], though she did not report the parameters used in the metric. Classification in the histogram model was made by calculating the likelihood of the target color triple for each color category, and returning a set of category-confidence pairs, sorted by likelihood.

Context comparator, S_c

Context similarity between two contexts was equal to the fraction of intersecting items plus a small constant.

$$S_c(c_i, c_j) = \epsilon + \frac{|c_i \cap c_j|}{|c_i \cup c_j|}$$

[1]The effective number of bins is less because some bins cover CIE L*a*b* values outside of the gamut of sRGB. Those colors appear nowhere in the corpus.

[2]I use "count" in quotations here because for this modified histogram, the total in all the bins no longer represents the number of examples encountered.

In the evaluation, each context pair had the same set of keys, so the above scoring metric is equivalent to

$$S_c(c_i, c_j) = \epsilon + |c_i \cap c_j|$$

The small constant was necessary in order to prevent similarity scores of zero from ever occurring. Similarity scores were used to weight the contribution of a stub classifier's categorization during a wake-cycle classification (see below), so scores of zero would nullify the effect of the classifier.

Stub-classifier comparator, S_f

Histogram classifiers were compared against each other using a distance metric $\Delta F(f_i^x, f_j^x)$ where f_i^x is the histogram for category x in classifier f_i. The final classifier similarity was then computed as

$$S_f(f_i, f_j) = 1 - \frac{1}{|c|} \sum_{x \in c} \Delta F(f_i^x, f_j^x)$$

Here, c represents the set of categories into which f_i and f_j can classify a feature vector.

Context Merging, $M(c_i, c_j)$

Two contexts were merged by taking their intersection: keys and values common to both contexts are included in the merged context; other are discarded.

For example:

$$M\left(\left\{\begin{array}{rl} id & : 10 \\ sex & : male \\ image & : flamingo \\ language & : english \end{array}\right\}, \left\{\begin{array}{rl} id & : 23 \\ sex & : female \\ image & : flamingo \\ language & : english \end{array}\right\}\right) = \left\{\begin{array}{rl} image & : flamingo \\ language & : english \end{array}\right\}$$

5.2 Procedure

Performance was evaluated using five-fold cross-validation which partitioned the training and test data by participant. The more detailed parameter space in Figure 5.2 was evaluated with leave-one-out cross-validation. A separate test set was not held out because the number of participants was small. Training proceeded in two phases. In the first phase, color categories were primed using the context-free class-membership data gathered in the experiment (Section 4.1.4). This training occurred in a single wake-phase. The second training phase used the context-sensitive data from the binary forced-choice task. Each wake phase in this training set contained the ten data points gathered for each color pair, context, and participant.

The stub-classifier similarity threshold γ was varied through a range of [0.75,1.0]. The reweighting parameter ω was varied from one to seventy.

Classifier accuracy was measured against the binary forced-choice data of the held-out participants. Recall that each data point represents an ambiguous color and the participant's category choice between two candidates. A classification was marked correct if the classifier ranked the participant's choice higher than the alternative color category.

5.3 Comparison Classifier

The baseline competitor compared against Context Dreaming was a histogram-classifier with no incorporation of context. Context Dreaming should be evaluated against the same type of online classifier used as its stub for fair comparison. Results reported for the competitor were gathered by the same procedure as for Context Dreaming.

5.4 Results and Discussion

The baseline histogram classifier correctly predicted a participant's responses 62.3% of the time. The peak score for the implemented Context Dreaming was 66.7% ($\omega = 3, \gamma = 0.99$), representing a 7.1% increase. Figures 5.1 and

5.2 show the effects of ω and γ on classifier performance and library size.

Overall, classification results for Context Dreaming were disappointing. As the figures indicate, the algorithm is highly sensitive to the ω and γ parameters, whose meaning, in turn, is entirely dependent on the implemented context data-structure and the two comparator functions. With the current implementation, the performance of the classifier drops below the baseline outside a narrow window of parameter values. In the course of implementing and testing the Context Dreaming framework, I learned that the choice of context representation and comparator function S_c are of critical importance for performance. The implemented context data-type does not balance the number of possible values each key can take. After a few dialogs, intersected contexts would favor keys like "sex", which could take only two possible values. Using this simple key-value mapping, there is no way to score the *relevance* of any key. The importance of "sex" or "id" is indistinguishable from "image" or "text" context keys. I believe that scoring classifiers at the end of a wake cycle as was discussed in Section 3.7.1, and augmenting key-value contexts with a key-relevance score, would be able to mitigate this problem and allow the classifier to perform better with a smaller library size.

The order in which a Context Dreaming classifier has dialogs and receives training data will also greatly affect performance. Early training data has a particularly pronounced effect. For this reason I believe the classifier performance as a function of γ and ω was noisy, even with leave-one-out cross-validation.

With the color experiment, I attempted to isolate an evaluation dataset where contextual effects are simple in their representation yet large in their magnitude. Unfortunately, I don't believe the dataset captures the richness of human experience that comes to bear on the color-naming problem; in fact, it may mask critical dimensions. Humans are always processing in context—we see this in the experimental data with the nearly-identical decision times of the forced-choice task. With only five color samples per combination of color pair, context and participant, the evaluation dataset is sparse which may also be a factor in the algorithm's disappointing performance.

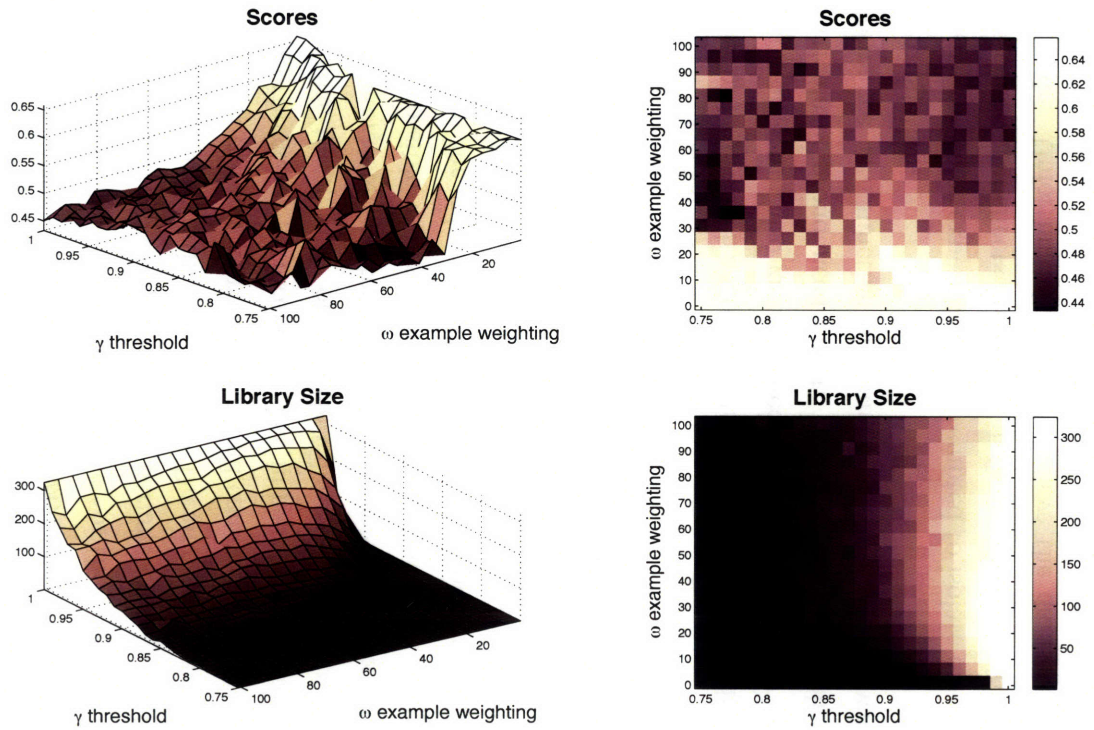


Figure 5.1: The effects of γ and ω on Context Dreaming performance and library size.

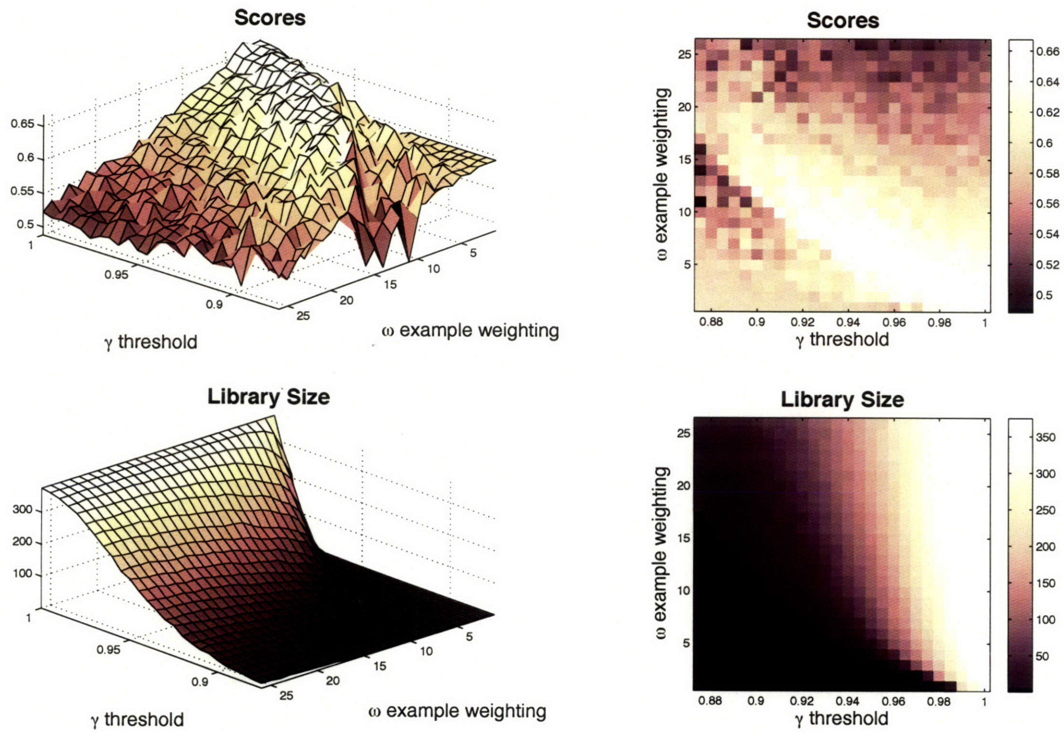


Figure 5.2: A more detailed look at the peak ranges of γ and ω. These results were generated using leave-one-out cross-validation.

CHAPTER 6

CONCLUSION

Labeling our world is almost never as simple as finding the word in the dictionary which matches the definition of what's being labeled. The world is noisy, and the partitions of labels have strange boundaries. Putting an intellectual framework on this gross task of categorization follows two paths. From the ground up, we have attempted to build machines that can distinguish the metaphorical wheat from the chaff and split the world into meaningful categories; and from the head down, the cognitive sciences have attempted to elucidate the mechanisms within us that make categorization seem like such an effortless task. This thesis attempts to add a small amount of knowledge to both of these camps.

The primary contribution of this thesis is the Context Dreaming algorithm, a classification mechanism bound to the real-life circumstances of finding shared meaning in conversation. Context Dreaming is a framework, a means with which classifiers of many sorts can gain sensitivity to context. This classifier rapidly adapts to the context in which it's asked to make classifications. Moreover, the information it gains during this quick adaptation is incorporated for future use. Context Dreaming is a straightforward algorithm, adding little computational complexity. Finally, Context Dreaming is a flexible system, fodder which gives many future avenues to explore.

The other half of the balance I hope to strike with this work is an experiment of human behaviour. Seeking to find a corpus on which I could evaluate the context-sensitive classifier, I developed and executed a study designed to quantify the some of the effects of semantic context on human categorization of colors. By presenting ambiguous colors on the border between two major color categories, then forcing a categorization between those color candidates, I was able to amplify the context effect. The results were clear: semantic context has a strong effect on color categorization. But the results were also surprising: some pairs of colors had an effect in the *opposite* direction as was hypothesized. This result led to a verifiable model which would explain the particular results.

Future Directions

Work on the ideas presented in this thesis has not come to an end, but rather to a moment of pause, reflection and summary. In both the cognitive science and computer science components of this work, there are elements which I hope to refine. Regarding the implemented Context Dreaming system: I hope to improve performance, both on synthetic and experimental data. Despite the disappointing performance on the color-context corpus, I believe that the core ideas of the algorithm are sound, and that further refinements—perhaps just those mentioned in the chapter describing the algorithm—will show my intuition to be fitting. Finally, to show that the concepts behind the Context Dreaming algorithm are durable, the refined implementation will need to be evaluated against a diverse collection of datasets.

Balancing the refinements of the algorithm I hope to complete, the surprises discovered in the context color-categorization experiment warrant further investigation. Specifically, I believe that the two experiments proposed in Chapter 4 will validate the hypothesis that the set of colors prototypical of a context will skew a participant's classification in the way observed in the collected data.

In closing...

Words bend their meaning to situation and to the person using them. Context matters. And perhaps not more so than in humor. What is funny? Image building a joke classifier.[1] Would building such a classifier even begin to be possible without taking context into account? Is a joke about subsumption architecture funny when the context is an NFL stadium locker room? I'm hardly making the claim that the Context Dreaming algorithm brings us materially closer to an automated joke classifier, but rather that context should be a critical constituent for many categorization problem-solvers.

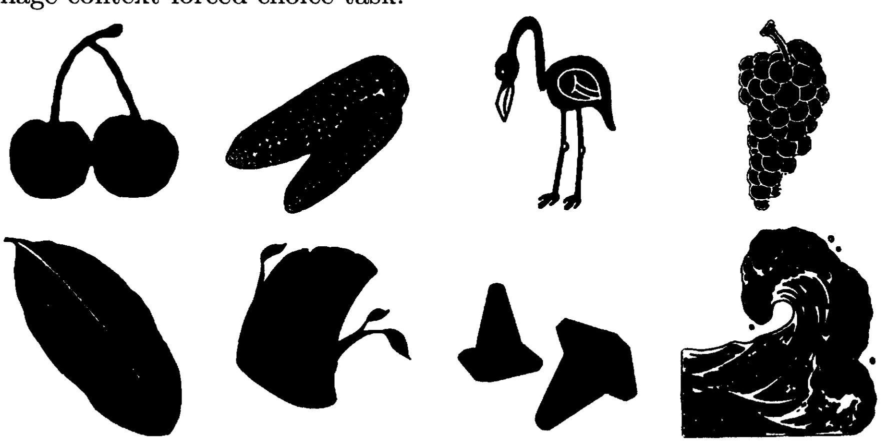
We should start with a taxi-driver's knowledge of colors. Then maybe move to a painter's knowledge. Jokes? Jokes will come later.

[1] Just such a classifier was imagined by Robert Heinlein in *The Moon is a Harsh Mistress*. Early in the novel, a computer tries to classify between "funny once" and "funny always" (page 17). The humor classification task finds many homes in science fiction. See also the beloved episode of *Star Trek: The Next Generation*, "The Outrageous Okona", where the android character Data seeks to understand the meaning of humor.

APPENDIX A

IMAGE STIMULUS USED IN THE EXPERIMENT

The following image stencils were used when displaying ambiguous colors in the image-context forced-choice task.



APPENDIX B

RESULTS: COLOR SURVEY TASK

The following pages contain the results of the color-survey task from the experiment described in chapter 4. Each number represents the count of participant which selected that patch either as a focal color (the best example of a color term) or as a class member (in the set covered by the color term). The array of color swatches matched the Munsell array of the World Color Survey, though constrained within the bounds of the sRGB gamut. There were a total of twenty-three participants.

The following table includes the CIE L*a*b* and sRGB values used for each swatch. RGB values are normalized to [0,1]. x and y are the coordinates of the swatch in the stimulus.

x	y	L*	a*	b*	R	G	B	Clipped
0	0	96.00	-0.06	0.06	0.954	0.955	0.954	
0	1	91.08	-0.05	0.06	0.899	0.900	0.899	
1	1	91.08	5.53	2.22	0.950	0.885	0.884	
2	1	91.08	5.51	3.28	0.953	0.885	0.876	
3	1	91.08	5.54	4.46	0.957	0.885	0.867	
4	1	91.08	5.43	5.64	0.960	0.885	0.858	
5	1	91.08	5.21	7.67	0.964	0.885	0.843	
6	1	91.08	4.30	10.08	0.965	0.886	0.825	
7	1	91.08	3.14	12.37	0.962	0.889	0.808	
8	1	91.08	1.28	14.41	0.954	0.893	0.792	
9	1	91.08	-0.46	29.79	0.980	0.893	0.676	
10	1	91.08	-5.25	45.24	0.977	0.902	0.556	

x	y	L*	a*	b*	R	G	B	Clipped
11	1	91.08	-9.03	45.94	0.952	0.910	0.549	
12	1	91.08	-12.17	45.90	0.929	0.917	0.549	
13	1	91.08	-16.65	44.66	0.894	0.927	0.558	
14	1	91.08	-14.67	29.61	0.874	0.926	0.675	
15	1	91.08	-10.76	13.38	0.858	0.921	0.798	
16	1	91.08	-12.29	10.56	0.836	0.925	0.819	
17	1	91.08	-13.13	7.63	0.819	0.927	0.841	
18	1	91.08	-13.24	5.48	0.810	0.928	0.857	
19	1	91.08	-13.22	3.82	0.804	0.928	0.869	
20	1	91.08	-12.96	2.42	0.801	0.928	0.880	
21	1	91.08	-12.69	1.23	0.798	0.928	0.889	
22	1	91.08	-11.94	-0.27	0.799	0.927	0.900	
23	1	91.08	-10.86	-2.13	0.800	0.925	0.914	
24	1	91.08	-9.69	-3.22	0.806	0.923	0.922	
25	1	91.08	-7.96	-4.41	0.816	0.919	0.931	
26	1	91.08	-6.32	-5.08	0.828	0.916	0.937	
27	1	91.08	-4.65	-5.34	0.841	0.912	0.939	
28	1	91.08	-3.48	-5.55	0.850	0.909	0.940	
29	1	91.08	-1.49	-5.61	0.867	0.905	0.941	
30	1	91.08	-0.37	-5.58	0.876	0.902	0.941	
31	1	91.08	1.12	-5.38	0.889	0.899	0.940	
32	1	91.08	2.19	-4.93	0.899	0.896	0.936	
33	1	91.08	2.95	-4.85	0.906	0.894	0.936	
34	1	91.08	3.36	-4.24	0.911	0.893	0.932	
35	1	91.08	4.31	-2.82	0.924	0.890	0.921	
36	1	91.08	4.70	-2.00	0.930	0.889	0.915	
37	1	91.08	5.04	-1.14	0.935	0.888	0.909	
38	1	91.08	5.27	-0.10	0.940	0.887	0.901	
39	1	91.08	5.41	0.74	0.944	0.886	0.895	
40	1	91.08	5.44	1.50	0.947	0.886	0.889	
0	2	81.35	-0.05	0.06	0.792	0.793	0.792	
1	2	81.35	23.67	9.47	0.986	0.728	0.727	
2	2	81.35	23.18	13.11	0.991	0.729	0.701	
3	2	81.35	22.67	17.48	0.997	0.729	0.669	
4	2	81.35	21.06	22.40	0.996	0.732	0.633	
5	2	81.35	18.52	26.72	0.988	0.739	0.601	
6	2	81.35	14.87	30.96	0.972	0.748	0.569	
7	2	81.35	14.98	47.04	0.998	0.744	0.446	

x	y	L*	a*	b*	R	G	B	Clipped
8	2	81.35	15.10	90.21	1.000	0.739	-0.000	*
9	2	81.35	7.28	109.12	0.998	0.759	-0.000	*
10	2	81.35	-3.69	100.18	0.928	0.786	-0.000	*
11	2	81.35	-10.94	87.90	0.875	0.803	-0.000	*
12	2	81.35	-17.25	87.99	0.833	0.816	-0.000	*
13	2	81.35	-26.04	85.63	0.771	0.834	-0.000	*
14	2	81.35	-30.49	65.57	0.716	0.844	0.268	
15	2	81.35	-40.58	54.46	0.613	0.863	0.369	
16	2	81.35	-38.45	34.02	0.578	0.862	0.535	
17	2	81.35	-42.41	23.78	0.497	0.870	0.612	
18	2	81.35	-33.80	12.41	0.539	0.858	0.697	
19	2	81.35	-33.63	8.67	0.522	0.859	0.725	
20	2	81.35	-33.04	4.96	0.508	0.858	0.752	
21	2	81.35	-32.18	1.81	0.500	0.858	0.775	
22	2	81.35	-21.19	-1.35	0.602	0.838	0.800	
23	2	81.35	-19.64	-4.30	0.603	0.836	0.821	
24	2	81.35	-18.27	-7.01	0.603	0.834	0.841	
25	2	81.35	-16.44	-9.33	0.609	0.831	0.858	
26	2	81.35	-13.22	-11.59	0.629	0.825	0.875	
27	2	81.35	-10.44	-12.79	0.650	0.819	0.884	
28	2	81.35	-6.82	-13.92	0.678	0.812	0.893	
29	2	81.35	-5.05	-23.28	0.645	0.811	0.961	
30	2	81.35	-0.24	-24.03	0.687	0.801	0.967	
31	2	81.35	3.22	-14.27	0.764	0.789	0.896	
32	2	81.35	5.82	-13.42	0.789	0.783	0.891	
33	2	81.35	8.19	-12.62	0.811	0.777	0.885	
34	2	81.35	9.68	-11.26	0.827	0.772	0.875	
35	2	81.35	18.74	-12.74	0.890	0.749	0.887	
36	2	81.35	21.03	-9.69	0.916	0.742	0.865	
37	2	81.35	22.62	-5.51	0.940	0.736	0.835	
38	2	81.35	23.53	-0.96	0.959	0.732	0.803	
39	2	81.35	23.82	2.52	0.970	0.730	0.778	
40	2	81.35	23.74	6.18	0.979	0.729	0.751	
0	3	71.60	-0.04	0.05	0.688	0.688	0.687	
1	3	71.60	32.18	12.57	0.932	0.598	0.604	
2	3	71.60	31.70	17.23	0.938	0.598	0.571	
3	3	71.60	38.22	29.22	0.994	0.573	0.487	
4	3	71.60	34.96	37.54	0.986	0.583	0.426	

x	y	L*	a*	b*	R	G	B	Clipped
5	3	71.60	30.67	44.56	0.970	0.595	0.373	
6	3	71.60	32.89	77.02	1.000	0.584	-0.000	*
7	3	71.60	24.12	84.39	0.962	0.611	-0.000	*
8	3	71.60	15.62	90.72	0.917	0.635	-0.000	*
9	3	71.60	6.22	81.95	0.859	0.660	-0.000	*
10	3	71.60	-2.30	85.64	0.810	0.680	-0.000	*
11	3	71.60	-9.51	87.14	0.767	0.696	-0.000	*
12	3	71.60	-15.98	87.62	0.726	0.709	-0.000	*
13	3	71.60	-24.87	85.15	0.665	0.726	-0.000	*
14	3	71.60	-34.49	79.15	0.591	0.743	-0.000	*
15	3	71.60	-40.28	53.23	0.507	0.755	0.278	
16	3	71.60	-46.70	41.31	0.411	0.766	0.377	
17	3	71.60	-51.55	27.69	0.282	0.775	0.481	
18	3	71.60	-41.95	14.43	0.344	0.763	0.579	
19	3	71.60	-41.85	9.77	0.312	0.764	0.613	
20	3	71.60	-41.54	5.41	0.280	0.765	0.644	
21	3	71.60	-40.91	1.15	0.247	0.765	0.675	
22	3	71.60	-39.21	-4.56	0.212	0.763	0.716	
23	3	71.60	-27.52	-7.11	0.391	0.745	0.735	
24	3	71.60	-25.13	-11.03	0.393	0.741	0.763	
25	3	71.60	-22.32	-14.67	0.402	0.737	0.789	
26	3	71.60	-18.37	-18.07	0.426	0.731	0.814	
27	3	71.60	-14.22	-19.90	0.463	0.724	0.827	
28	3	71.60	-12.05	-29.46	0.413	0.723	0.896	
29	3	71.60	-5.09	-30.75	0.489	0.709	0.905	
30	3	71.60	0.85	-31.32	0.549	0.696	0.910	
31	3	71.60	6.06	-22.71	0.645	0.681	0.849	
32	3	71.60	10.25	-21.60	0.685	0.670	0.842	
33	3	71.60	13.67	-19.66	0.720	0.661	0.828	
34	3	71.60	16.13	-17.52	0.746	0.654	0.813	
35	3	71.60	26.62	-18.32	0.819	0.625	0.820	
36	3	71.60	29.26	-13.95	0.850	0.616	0.790	
37	3	71.60	38.97	-10.98	0.921	0.583	0.770	
38	3	71.60	40.04	-2.70	0.948	0.576	0.712	
39	3	71.60	32.47	2.76	0.914	0.600	0.673	
40	3	71.60	32.41	7.97	0.925	0.598	0.636	
0	4	61.70	-0.04	0.04	0.584	0.584	0.584	
1	4	61.70	49.42	18.23	0.922	0.429	0.468	

x	y	L*	a*	b*	R	G	B	Clipped
2	4	61.70	48.53	25.92	0.927	0.431	0.416	
3	4	61.70	46.38	35.47	0.926	0.437	0.349	
4	4	61.70	49.15	56.82	0.956	0.422	0.185	
5	4	61.70	44.69	79.79	0.943	0.439	-0.000	*
6	4	61.70	29.38	64.40	0.858	0.495	0.083	
7	4	61.70	22.09	71.31	0.822	0.517	-0.000	*
8	4	61.70	14.45	76.93	0.783	0.538	-0.000	*
9	4	61.70	5.87	68.59	0.732	0.560	-0.000	*
10	4	61.70	-1.40	71.61	0.691	0.576	-0.000	*
11	4	61.70	-7.92	73.21	0.653	0.590	-0.000	*
12	4	61.70	-14.02	73.17	0.616	0.603	-0.000	*
13	4	61.70	-21.81	70.54	0.563	0.617	-0.000	*
14	4	61.70	-30.13	65.11	0.499	0.632	-0.000	*
15	4	61.70	-46.97	64.65	0.355	0.657	-0.000	*
16	4	61.70	-55.94	49.19	0.195	0.671	0.208	
17	4	61.70	-51.58	26.30	0.122	0.668	0.393	
18	4	61.70	-52.22	16.26	-0.000	0.670	0.465	*
19	4	61.70	-52.09	10.63	-0.000	0.671	0.504	*
20	4	61.70	-51.88	5.51	-0.000	0.672	0.540	*
21	4	61.70	-41.11	0.19	0.006	0.658	0.578	
22	4	61.70	-39.17	-5.88	-0.000	0.657	0.621	*
23	4	61.70	-36.77	-10.90	-0.000	0.654	0.656	*
24	4	61.70	-33.16	-16.42	-0.000	0.650	0.694	*
25	4	61.70	-29.22	-21.27	-0.000	0.646	0.728	*
26	4	61.70	-22.98	-25.73	0.097	0.637	0.760	
27	4	61.70	-16.81	-28.31	0.224	0.627	0.778	
28	4	61.70	-13.07	-37.55	0.139	0.623	0.843	
29	4	61.70	-4.52	-39.18	0.303	0.607	0.855	
30	4	61.70	2.64	-39.49	0.398	0.592	0.857	
31	4	61.70	9.58	-30.64	0.530	0.573	0.796	
32	4	61.70	15.28	-29.05	0.586	0.558	0.786	
33	4	61.70	20.07	-26.92	0.632	0.545	0.772	
34	4	61.70	23.29	-23.79	0.668	0.534	0.750	
35	4	61.70	34.47	-24.03	0.743	0.501	0.753	
36	4	61.70	37.10	-18.34	0.778	0.489	0.714	
37	4	61.70	38.67	-11.63	0.806	0.481	0.669	
38	4	61.70	40.07	-3.44	0.833	0.473	0.613	
39	4	61.70	49.38	2.78	0.898	0.434	0.573	

x	y	L*	a*	b*	R	G	B	Clipped
40	4	61.70	49.67	10.71	0.913	0.430	0.519	
0	5	51.57	-0.03	0.04	0.482	0.482	0.482	
1	5	51.57	59.36	19.67	0.852	0.268	0.363	
2	5	51.57	58.01	30.52	0.856	0.273	0.292	
3	5	51.57	55.76	42.05	0.853	0.283	0.211	
4	5	51.57	55.20	68.32	0.861	0.283	-0.000	*
5	5	51.57	40.53	69.04	0.791	0.353	-0.000	*
6	5	51.57	29.45	64.44	0.735	0.394	-0.000	*
7	5	51.57	19.79	58.75	0.683	0.424	0.002	
8	5	51.57	13.12	63.99	0.651	0.441	-0.000	*
9	5	51.57	5.33	54.90	0.604	0.461	0.054	
10	5	51.57	-0.48	57.27	0.573	0.474	-0.000	*
11	5	51.57	-6.40	58.71	0.540	0.486	-0.000	*
12	5	51.57	-11.71	58.46	0.508	0.497	-0.000	*
13	5	51.57	-18.37	55.88	0.463	0.509	0.004	
14	5	51.57	-25.02	50.55	0.412	0.520	0.093	
15	5	51.57	-40.14	52.84	0.291	0.543	0.045	
16	5	51.57	-56.86	49.22	-0.000	0.564	0.093	*
17	5	51.57	-63.28	28.95	-0.000	0.572	0.274	*
18	5	51.57	-52.87	15.46	-0.000	0.563	0.372	*
19	5	51.57	-52.69	9.70	-0.000	0.564	0.411	*
20	5	51.57	-51.99	4.46	-0.000	0.564	0.447	*
21	5	51.57	-51.20	-1.36	-0.000	0.564	0.486	*
22	5	51.57	-38.59	-7.00	-0.000	0.549	0.525	*
23	5	51.57	-36.19	-11.99	-0.000	0.547	0.558	*
24	5	51.57	-32.36	-17.71	-0.000	0.544	0.597	*
25	5	51.57	-27.95	-22.24	-0.000	0.538	0.628	*
26	5	51.57	-21.72	-26.63	-0.000	0.530	0.657	*
27	5	51.57	-15.72	-29.08	0.029	0.521	0.674	
28	5	51.57	-11.88	-38.56	-0.000	0.518	0.739	*
29	5	51.57	-3.41	-48.08	-0.000	0.507	0.804	*
30	5	51.57	5.02	-48.35	0.200	0.490	0.807	
31	5	51.57	14.19	-39.14	0.413	0.464	0.745	
32	5	51.57	22.05	-37.02	0.494	0.442	0.731	
33	5	51.57	27.40	-33.72	0.550	0.426	0.709	
34	5	51.57	31.24	-29.94	0.592	0.412	0.684	
35	5	51.57	35.21	-24.66	0.636	0.396	0.649	
36	5	51.57	46.17	-23.40	0.706	0.354	0.642	

x	y	L*	a*	b*	R	G	B	Clipped
37	5	51.57	48.98	-15.95	0.741	0.338	0.593	
38	5	51.57	51.20	-5.84	0.775	0.323	0.527	
39	5	51.57	59.69	1.53	0.831	0.274	0.481	
40	5	51.57	59.79	10.72	0.844	0.269	0.421	
0	6	41.22	-0.03	0.03	0.381	0.381	0.381	
1	6	41.22	61.40	17.92	0.735	0.116	0.280	
2	6	41.22	59.50	30.17	0.736	0.131	0.202	
3	6	41.22	56.60	40.99	0.728	0.155	0.128	
4	6	41.22	51.06	58.81	0.709	0.193	-0.000	*
5	6	41.22	32.37	46.62	0.619	0.285	0.061	
6	6	41.22	22.15	41.13	0.567	0.319	0.107	
7	6	41.22	17.04	45.95	0.544	0.333	0.054	
8	6	41.22	9.31	37.80	0.499	0.354	0.129	
9	6	41.22	4.42	40.43	0.475	0.365	0.104	
10	6	41.22	-0.21	42.37	0.452	0.375	0.082	
11	6	41.22	-4.87	43.26	0.426	0.384	0.070	
12	6	41.22	-9.31	43.05	0.400	0.392	0.069	
13	6	41.22	-14.57	40.52	0.366	0.402	0.094	
14	6	41.22	-19.48	36.56	0.328	0.410	0.128	
15	6	41.22	-33.18	41.15	0.227	0.430	0.078	
16	6	41.22	-38.39	31.53	0.141	0.437	0.164	
17	6	41.22	-53.57	23.28	-0.000	0.455	0.221	*
18	6	41.22	-54.08	14.27	-0.000	0.456	0.283	*
19	6	41.22	-53.43	8.61	-0.000	0.457	0.320	*
20	6	41.22	-52.68	3.36	-0.000	0.457	0.354	*
21	6	41.22	-41.21	-2.07	-0.000	0.446	0.390	*
22	6	41.22	-39.09	-8.33	-0.000	0.444	0.431	*
23	6	41.22	-36.33	-13.61	-0.000	0.442	0.465	*
24	6	41.22	-24.01	-14.24	-0.000	0.426	0.470	*
25	6	41.22	-20.46	-17.64	-0.000	0.422	0.492	*
26	6	41.22	-21.08	-28.21	-0.000	0.426	0.560	*
27	6	41.22	-14.67	-30.69	-0.000	0.417	0.577	*
28	6	41.22	-10.54	-39.65	-0.000	0.414	0.636	*
29	6	41.22	-2.09	-40.81	-0.000	0.400	0.644	*
30	6	41.22	7.17	-48.94	0.003	0.385	0.698	
31	6	41.22	20.81	-48.05	0.299	0.353	0.693	
32	6	41.22	24.29	-38.02	0.399	0.337	0.627	
33	6	41.22	29.49	-34.06	0.456	0.319	0.602	

x	y	L*	a*	b*	R	G	B	Clipped
34	6	41.22	33.08	-30.50	0.494	0.306	0.579	
35	6	41.22	36.48	-25.75	0.531	0.291	0.549	
36	6	41.22	46.93	-24.70	0.594	0.246	0.543	
37	6	41.22	49.84	-17.34	0.628	0.226	0.497	
38	6	41.22	51.84	-8.08	0.656	0.209	0.439	
39	6	41.22	52.58	-0.05	0.673	0.200	0.389	
40	6	41.22	61.57	8.63	0.727	0.121	0.337	
0	7	30.77	-0.02	0.03	0.284	0.284	0.284	
1	7	30.77	47.34	12.58	0.545	0.106	0.217	
2	7	30.77	46.08	20.60	0.546	0.113	0.169	
3	7	30.77	52.68	34.06	0.582	0.031	0.087	
4	7	30.77	40.89	36.03	0.532	0.145	0.065	
5	7	30.77	28.16	36.05	0.476	0.203	0.056	
6	7	30.77	18.18	30.27	0.426	0.236	0.097	
7	7	30.77	13.89	33.80	0.408	0.247	0.067	
8	7	30.77	9.51	37.16	0.389	0.258	0.031	
9	7	30.77	3.10	25.71	0.348	0.273	0.124	
10	7	30.77	-0.15	26.58	0.332	0.280	0.117	
11	7	30.77	-3.32	26.87	0.315	0.286	0.114	
12	7	30.77	-6.44	26.57	0.297	0.291	0.115	
13	7	30.77	-10.05	24.97	0.273	0.298	0.125	
14	7	30.77	-13.04	21.97	0.250	0.303	0.145	
15	7	30.77	-23.96	27.29	0.179	0.318	0.105	
16	7	30.77	-27.26	20.69	0.129	0.323	0.151	
17	7	30.77	-39.23	16.60	-0.000	0.337	0.176	*
18	7	30.77	-39.62	10.61	-0.000	0.338	0.215	*
19	7	30.77	-49.93	6.94	-0.000	0.348	0.237	*
20	7	30.77	-39.18	2.14	-0.000	0.339	0.267	*
21	7	30.77	-29.07	-1.92	-0.000	0.329	0.293	*
22	7	30.77	-27.32	-6.72	-0.000	0.328	0.322	*
23	7	30.77	-25.67	-10.89	-0.000	0.327	0.348	*
24	7	30.77	-22.84	-15.29	-0.000	0.324	0.375	*
25	7	30.77	-19.73	-18.89	-0.000	0.321	0.397	*
26	7	30.77	-15.41	-22.00	-0.000	0.316	0.417	*
27	7	30.77	-10.88	-24.20	-0.000	0.310	0.431	*
28	7	30.77	-7.86	-33.80	-0.000	0.309	0.490	*
29	7	30.77	0.29	-43.03	-0.000	0.300	0.549	*
30	7	30.77	7.86	-43.39	-0.000	0.285	0.552	*

x	y	L*	a*	b*	R	G	B	Clipped
31	7	30.77	25.75	-49.63	0.217	0.245	0.592	
32	7	30.77	28.36	-39.75	0.314	0.230	0.530	
33	7	30.77	33.51	-35.69	0.370	0.209	0.505	
34	7	30.77	36.90	-32.14	0.405	0.193	0.483	
35	7	30.77	39.78	-27.99	0.435	0.178	0.458	
36	7	30.77	42.44	-23.60	0.462	0.161	0.431	
37	7	30.77	44.87	-16.65	0.491	0.142	0.389	
38	7	30.77	46.25	-9.51	0.511	0.128	0.347	
39	7	30.77	46.93	-2.51	0.526	0.118	0.305	
40	7	30.77	47.63	5.24	0.539	0.108	0.260	
0	8	20.54	-0.02	0.02	0.194	0.194	0.194	
1	8	20.54	36.82	3.13	0.378	0.068	0.181	
2	8	20.54	36.64	9.24	0.384	0.066	0.148	
3	8	20.54	35.15	15.44	0.382	0.075	0.113	
4	8	20.54	25.15	16.60	0.341	0.125	0.103	
5	8	20.54	15.99	14.49	0.299	0.155	0.114	
6	8	20.54	13.28	18.21	0.290	0.162	0.091	
7	8	20.54	10.16	21.44	0.279	0.169	0.070	
8	8	20.54	3.62	11.09	0.236	0.185	0.131	
9	8	20.54	1.58	11.89	0.227	0.189	0.126	
10	8	20.54	-0.11	12.03	0.218	0.192	0.125	
11	8	20.54	-1.74	11.57	0.209	0.195	0.127	
12	8	20.54	-3.09	10.75	0.200	0.197	0.132	
13	8	20.54	-4.47	9.36	0.189	0.200	0.140	
14	8	20.54	-5.36	7.98	0.181	0.202	0.148	
15	8	20.54	-14.85	15.25	0.134	0.215	0.103	
16	8	20.54	-16.46	11.28	0.110	0.217	0.127	
17	8	20.54	-17.35	7.46	0.087	0.219	0.149	
18	8	20.54	-27.79	7.33	-0.000	0.231	0.149	*
19	8	20.54	-27.43	4.26	-0.000	0.231	0.167	*
20	8	20.54	-26.87	1.22	-0.000	0.231	0.185	*
21	8	20.54	-17.18	-1.47	0.024	0.220	0.201	
22	8	20.54	-16.13	-4.67	0.007	0.220	0.219	
23	8	20.54	-15.07	-7.11	-0.000	0.219	0.233	*
24	8	20.54	-13.39	-9.92	-0.000	0.217	0.249	*
25	8	20.54	-11.72	-12.03	-0.000	0.216	0.262	*
26	8	20.54	-9.52	-13.99	0.011	0.213	0.273	
27	8	20.54	-9.22	-24.04	-0.000	0.216	0.332	*

x	y	L*	a*	b*	R	G	B	Clipped
28	8	20.54	-5.25	-25.49	-0.000	0.211	0.340	*
29	8	20.54	0.11	-26.61	-0.000	0.202	0.347	*
30	8	20.54	7.63	-35.92	-0.000	0.193	0.403	*
31	8	20.54	23.81	-41.87	0.154	0.157	0.439	
32	8	20.54	23.12	-31.33	0.221	0.152	0.376	
33	8	20.54	26.82	-27.84	0.261	0.137	0.356	
34	8	20.54	29.26	-25.01	0.285	0.126	0.340	
35	8	20.54	22.74	-16.65	0.276	0.144	0.291	
36	8	20.54	24.14	-14.33	0.290	0.139	0.278	
37	8	20.54	34.44	-14.69	0.340	0.094	0.281	
38	8	20.54	35.44	-10.40	0.353	0.086	0.257	
39	8	20.54	35.97	-6.33	0.362	0.080	0.234	
40	8	20.54	36.42	-2.08	0.370	0.074	0.210	
0	9	15.60	-0.02	0.02	0.153	0.153	0.153	

Focal: black

0																																								
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
23																																								

Focal: red

0																																								
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	1	1	15	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	1	0	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0																																								

Focal: purple

0																																								
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	3	0	0	0	1	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	3	5	4	2	0	1	0	1
0																																								

Focal: blue

0																																								
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	2	1	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	3	4	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	1	0	0	0	0	0	0	0	0	0
0																																								

Focal: brown

0																																								
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	1	2	9	7	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0																																								

Focal: gray

0																																								
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0																																								

Focal: orange

```
0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 1 6 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 7 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 1 2 2 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0
```

Focal: white

```
22
0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0
```

Focal: green

```
0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 2 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 9 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 2 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 2 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0
```

Focal: yellow

0																																								
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	5	12	3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0																																								

Focal: pink

0																																								
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	2	2
0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5
0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0																																								

Class: black

0																																								
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
23																																								

Class: red

0																																								
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	1	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	3	12	14	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	8	19	21	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	7	12	14	5	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0																																								

Class: purple

0																																								
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	3	4	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	5	6	6	0	0	1	1	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	9	12	12	2	0	1	1	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	15	14	6	2	1	1	1	1
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	12	19	16	12	5	3	2	1	1
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	14	20	18	18	12	5	4	2	2
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11	20	20	19	19	9	7	5	3
0																																								

Class: blue

0																																								
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	2	2	2	2	2	2	2	2	2	1	1	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	3	4	5	8	11	14	13	14	14	15	15	12	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2	3	4	5	9	11	16	16	16	16	20	18	14	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	2	1	4	9	14	20	21	21	21	22	23	16	1	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	3	9	15	17	19	20	21	22	22	23	2	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	15	15	15	20	20	23	22	23	8	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	16	15	16	18	19	23	23	22	10	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	12	16	16	18	21	23	23	23	14	0	0	0	0	0	0	0	0	0
0																																								

Class: brown

0																																								
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	1	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	1	3	3	4	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	2	9	15	14	9	8	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	18	19	20	17	17	14	4	3	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0																																								

Class: gray

0																																								
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
22	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
22	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
23	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
22	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1																																								

Class: orange

0																																								
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	1	1	1	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	2	5	6	17	11	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	2	23	23	18	11	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	2	6	21	22	17	10	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	2	5	14	15	9	5	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	2	2	6	5	4	3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0																																								

Class: white

23																																								
19	7	8	8	7	7	5	5	5	0	0	0	0	0	0	3	5	5	6	3	4	6	6	6	7	6	7	7	7	5	6	5	4	4	3	5	4	4	5	4	5
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0																																								

Class: green

0																																								
0	0	0	0	0	0	0	0	0	0	1	1	1	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	1	1	6	12	20	16	10	4	3	3	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	1	2	6	13	20	18	13	6	3	2	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	1	2	9	15	22	22	15	7	5	2	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	1	1	2	10	14	22	23	18	9	4	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	1	2	2	8	13	22	22	18	10	5	3	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	1	1	5	10	20	19	17	11	9	6	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	17	17	13	12	10	5	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0																																								

Class: yellow

0																																								
0	0	0	0	0	0	0	0	0	9	16	14	10	7	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	5	22	22	20	14	6	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	1	1	5	5	11	15	17	8	5	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	1	1	4	4	7	6	6	4	4	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	1	2	2	0	1	3	3	2	2	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0																																								

Class: pink

```
0
0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  1  0  0
0 14 13 10  4  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  1  2  5  8  8  9
1 14 12  6  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  1  2 13 15 14 15
0 14 10  2  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  2  7 13 16 17
0 16  4  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  5 11 15 15
0 12  3  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  3  5  8 10
0  4  2  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  1  4  5
0  3  1  0  0  1  1  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  3
0
```

BIBLIOGRAPHY

[1] K. Barnard, V. Cardei, and B. Funt. A comparison of computational color constancy algorithms. I: Methodology and experiments with synthesized data. *IEEE Transactions on Image Processing*, 11(9):972–984, 2002.

[2] T. Belpaeme. Reaching coherent color categories through communication. In B. Kröse, G. Schreiber de Rijke, and M. van Someren, editors, *Proceedings of the 13th Belgium-Netherlands Conference on Artificial Intelligence (BNAIC'01)*, pages 41–48, Amsterdam, The Netherlands, 2001.

[3] B. Berlin and P. Kay. *Basic Color Terms: Their Universality and Evolution.* University of California Press, 1969.

[4] Nick Braisby and Julie Dockrell. Why is colour naming difficult? *Journal of Child Language*, 26(1):23–47, Feb 1999.

[5] Leo Breiman. Bagging predictors. *Machine Learning*, 24(2):123–140, 1996.

[6] Susan E. Brennan and Herbert H. Clark. Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 22(6):1482–1496, 1996.

[7] A. Chella, M. Frixione, and S. Gaglio. Conceptual Spaces for Computer Vision Representations. *Artificial Intelligence Review*, 16(2):137–152, 2001.

[8] Heng-Da Cheng, Xihua Jiang, Ying Sun, and Jingli Wang. Color image segmentation: advances and prospects. *Pattern Recognition*, 34(12):2259–2281, 2001.

[9] Herbert H. Clark and Deanna Wilkes-Gibbs. Referring as a collaborative process. *Cognition*, 22:1–39, 1986.

[10] Y. Freund and R.E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55(1):119–139, 1997.

[11] Peter Gärdenfors. *Conceptual Spaces: The Geometry of Thought.* MIT Press, Cambridge, MA, 2000.

[12] Peter Gärdenfors and Mary-Anne Williams. Reasoning about categories in conceptual spaces. In *Proceedings of the Fourteenth International Joint Conference of Artificial Intelligence*, pages 385–392. Morgan Kaufmann, 2001.

[13] Thorsten Hansen, Maria Olkkonen, Sebastian Walter, and Karl R Gegenfurtner. Memory modulates color appearance. *Nature Neuroscience*, 9(11):1367–1368, 2006.

[14] C.L. Hardin and Luisa Maffi, editors. *Color Categories in Thought and Language.* Cambridge University Press, 1997.

[15] Peter A. Heeman and Graeme Hirst. Collaborating on referring expressions. *Computational Linguistics*, 21(3):351–382, 1995.

[16] Ewald Hering. *Outlines of a Theory of the Light Sense.* Harvard University Press, 1964.

[17] Geoffrey E. Hinton, Peter Dayan, Brendan J. Frey, and Radford M. Neal. The wake-sleep algorithm for unsupervised neural networks. *Science*, 268:1158–1161, 1995.

[18] Daoyun Ji and Matthew A. Wilson. Coordinated memory replay in the visual cortex and hippocampus during sleep. *Nature Neuroscience*, 10(1):100–107, December 2006.

[19] Paul Kay, Brent Berlin, Luisa Maffi, and William R. Merrifield. *World Color Survey.* Center for the Study of Language and Information, July 2007 [2003].

[20] J. Kittler, M. Hatef, RPW Duin, and J. Matas. On combining classifiers. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 20(3):226–239, 1998.

[21] George Lakoff. Hedges: A study in meaning criteria and the logic of fuzzy concepts. *Journal of Philosophical Logic*, 2(4):458–808, October 1973.

[22] Johan M. Lammens. *A Computational Model of Color Perception and Color Naming*. PhD thesis, SUNY/Buffalo, 24, 1994.

[23] David L. MacAdam. Visual sensitivities to color differences in daylight. *Journal of the Optical Society of America*, 32(5):247–274, May 1942.

[24] A. Maerz and M. Rea Paul. *A dictionary of color*. McGraw-Hill Book Company, New York, 1950.

[25] James L. McClelland, Bruce McNaughton, and Randall C. O'Reilly. Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, 102(3):419–457, 1995.

[26] Chirag Mehta. Name that color javascript. http://chir.ag/tech/download/ntc/.

[27] P. Melville and R. Mooney. Constructing Diverse Classifier Ensembles Using Artificial Training Examples. *Proceedings of the IJCAI*, pages 505–510, 2003.

[28] G. Menegaz, A. Le Troter, J. Sequeira, and J. M. Boi. A discrete model for color naming. *EURASIP Journal on Advances in Signal Processing*, 2007.

[29] A. Mojsilovic. A computational model for color naming and describing color composition of images. *IEEE Transactions on Image Processing*, 14(5):690–699, 2005.

[30] Nathan Moroney, Mark D. Fairchild, Robert W. G. Hunt, Changjun Li, M. Ronnier Luo, and Todd Newman. The CIECAM02 Color Appearance Model. In *The Tenth Color Imaging Conference: Color Science and Engineering*, pages 23–27. IS&T - The Society for Imaging Science and Technology, 2002.

[31] D. Nickerson. OSA color scale samples: a unique set. *Color Research & Application*, 6:7–33, 1981.

[32] Tuuli Nurminen, Jyri Kivinen, and Pirkko Oittinen. Color naming and computational prediction from natural images. *Graphic Arts in Finland*, 35(2):13–28, 2006.

[33] Terry Regier, Paul Kay, and Naveen Khetarpal. Color naming reflects optimal partitions of color space. *PNAS*, 104(4):1436–1441, January 2007.

[34] Deb Roy. Grounding words in perception and action: computational insights. *Trends in Cognitive Science*, 9(8):389–396, August 2005.

[35] Bilge Sayim, Kimberly A. Jameson, Nancy Alvarado, and Monika K. Szeszel. Semantic and perceptual representations of color: Evidence of a shared color-naming function. *Journal of Cognition and Culture*, 5(3-4):427–486, 2005.

[36] Robert E. Schapire. The strength of weak learnability. *Machine Learning*, 5:197–227, 1990.

[37] Robert E. Schapire. The boosting approach to machine learning: An overview. In D. D. Denison, M. H. Hansen, C. Holmes, B. Mallick, and B. Yu, editors, *Nonlinear Estimation and Classification*. Springer, 2003.

[38] D. Song and P. D. Bruza. Towards context sensitive information inference. *Journal of the American Society for Information Science and Technology*, 54(4):321–334, February 2003.

[39] Luc Steels and Tony Belpaeme. Coordinating perceptually grounded categories through language: A case study for colour. *Behavioral and Brain Sciences*, 28(4):469–89, August 2005. Target Paper, discussion 489-529.

[40] J. R. Stroop. Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 12:643–662, 1935.

[41] Johann Wolfgang von Goethe. *Theory of Colours*. The MIT Press, 1810 (1970).

[42] D.H. Wolpert. Stacked generalization. *Neural Networks*, 5(2):241–259, 1992.

[43] Günther Wyszecki and W. S. Stiles. *Color Science: Concepts and Methods, Quantitative Data and Formulae.* Wiley-Interscience, second edition, 2000.