DeepColor: Reinforcement Learning optimizes information efficiency and well-formedness in color name partitioning

Citation for the original published paper (version of record):
Kågebäck, M., Dubhashi, D., Sayeed, A. (2018). DeepColor: Reinforcement Learning optimizes information efficiency and well-formedness in color
name partitioning. Proceedings of the 40th Annual Meeting of the Cognitive Science Society (CogSci): 1895-1900

DeepColor: Reinforcement Learning optimizes information efficiency and well-formedness in color name partitioning

Mikael Kågebäck, Devdatt Dubhashi (kageback, dubhashi@chalmers.se)
Dept. of Computer Science and Engineering
Chalmers University of Technology, Sweden

Asad Sayeed (asad.sayeed@gu.se)
Dept. of Philosophy, Linguistics, and Theory of Science
University of Gothenburg, Sweden

Abstract

As observed in the World Color Survey (WCS), some universal properties can be identified in color naming schemes over a large number of languages. For example, Regier, Kay, and Khetrapal (2007) and Regier, Kemp, and Kay (2015); Gibson et al. (2017) recently explained these universal patterns in terms of near optimal color partitions and information theoretic measures of efficiency of communication. Here, we introduce a computational learning framework with multi-agent systems trained by reinforcement learning to investigate these universal properties. We compare the results with Regier et al. (2007, 2015) and show that our model achieves excellent quantitative agreement. This work introduces a multi-agent reinforcement learning framework as a powerful and versatile tool to investigate such semantic universals in many domains and contribute significantly to central questions in cognitive science.

Keywords: color naming; world color survey; reinforcement learning

Introduction

Semantic partitioning of the color space in the lexicon

Color naming universals have a long history in linguistic research (Berlin & Kay, 1969). At an individual level, color perception is subjective; it differs for biological reasons across individuals (extreme examples being colorblindness and tetrachromacy). There are commonly-observed differences in individual color–naming choices. What is "turquoise" to one person may be a variant of "blue" to another. Nevertheless, within the same linguistic milieu, there is overall agreement as to color-naming; most English-speaking people recognise the typical supermarket tomato as "red".

Berlin and Kay showed across a survey of 20 languages that there are strong consistencies in color naming and produced a set of universals: e.g., there are a maximum of eleven major color categories and, where fewer than eleven are realized for a given language, there is a standard pattern of emergence. This work came under methodological criticism (Lucy, 1997; Saunders, 1995), particularly the use of standardized color systems to abstract away from the interactional and cultural basis of color identification.

Given this methodological conflict, is it really the case that such universals are artifacts of a non-interactional method of investigation? Accounting for patterns of wide but constrained variation that have been observed empirically is a central challenge in understanding why languages have the particular forms they do. There is an increasingly influential proposal that language is shaped by the need for efficient communication (Piantadosi, Tily, & Gibson, 2011), which by its nature involves a trade–off between simplicity in an information–theoretic sense, which minimizes cognitive load, and informativeness which maximizes communicative effectiveness. Specifically, that good systems of categories have a near–optimal trade–off between these constraints.

Examples formalized in information-theoretic terms include suggestions that word frequency distributions, syllable durations, word lengths, syntactic structures, and case marking all facilitate efficient communication. This type of proposal has been particularly successful in explaining aspects of the structure of utterances as they unfold over time: linguistic structure optimizes information transfer between interacting language users. A different challenge is explaining the partitioning of semantic spaces—the common recognition of lexical-semantic differences across a speech community—likewise in terms of interactive behavior.

Color terms represent a limited semantic domain with easily manipulated parameters. By gradual changes of color value, an experimenter can manipulate red into orange, unlike other semantic domains, where the distinctions between potential referents (e.g., "car" vs. "truck") are not easily captured in explicit terms. In addition, recent work (Regier et al., 2015; Gibson et al., 2017) argues that color categories in language should support efficient communication.

Our goal in this paper is to investigate the scientific question of the emergence of color terms via a computational framework for modeling the partitioning of semantic spaces. Our framework, DeepColor, is based on a deep reinforcement learning approach with independent justification as a cognitive modeling paradigm. The framework, inspired by Regier et al. (2015) and Gibson et al. (2017), takes the form of two agents, one of which is attempting to communicate a color to the other through a channel with a limited selection of signals intended to refer to the colors in a commonly-used color-identification experimentation system. Learning takes place through a reward system as the agents converge on approximate color-meanings for the signal vocabulary. We show that the simulations we run in this noisy-channel environment converge on characteristics similar to cross-language human-collected data using the same color system. This work validates an approach using deep reinforcement learning with communicating agents as a research paradigm for color space partitioning in language.

Communicating color

Developed in 1905, the Munsell color system uses three color dimensions (hue, value, and chroma) to represent colors based on an "equidistance" metric calibrated experimentally by Albert Munsell. The World Color Survey (WCS; e.g. figure 4) uses the Munsell color system in a matrix arranged by 40 hues, 8 values (lightness), and at maximum chroma (saturation). A color map can be developed for a particular language by asking speakers of that language to name each color. Color identification boundaries can be compared across languages using the WCS mapping.

The WCS color map technique enables the testing of automatic systems to partition colors. Regier et al. (2007) experiment with partitioning the color space using a distance metric as a clustering criterion. They find a good distance metric by translating the WCS color map to the CIELAB space. CIELAB enables the translation of the WCS colors to a three-dimensional space, wherein the WCS colors appear to take an irregular spherical form. Regier et al. then use a well-formedness metric based on a similarity/dissimilarity trade–off to automatically construct color partitions in the CIELAB space. Regier et al. find correspondences between optimal color partitions and observed color maps from human surveys as well as determine that rotating the WCS color space for a given observed color map causes reduced well-formedness in the corresponding CIELAB space. This is preliminary evidence for the optimality of color spaces in human language in relation to a well-formedness trade-off statistic.

Following their earlier work, Regier et al. adopt an information–theoretic approach (2015) by introducing a communication system between two agents for multiple semantic domains (including color) and the corresponding notion of reconstruction error as the relative entropy (Kullback–Leibler divergence). The relative entropy is computed between the speaker's model and the listener's model of the probability that a particular term encodes a particular color. This becomes the communicative cost of a color labeling system. Regier et al. (2015) then show that real-world color-naming systems not only tend to have high well-formedness, but they also have low communicative cost. A similar framework is adopted in Gibson et al. (2017).

Reinforcement Learning: a general cognitive mechanism

Reinforcement learning (RL) studies the way that natural and artificial systems can learn to predict the consequences of and optimize their behavior in environments in which actions lead them from one state or situation to the next, and can also lead to rewards and punishments (Sutton & Barto, 1998; Wiering & van Otterlo, 2012). Such environments arise in a wide range of fields, including ethology, economics, psychology, and control theory. RL, originally born out of mathematical psychology and operations research, provides qualitative and quantitative computational-level models of these solutions.

There is an increasing realization that RL may offer more than just a computational, theory for affective decision-making but that RL algorithms appear to be directly instantiated in neural mechanisms, such as the phasic activity of dopamine neurons (Dayan & Niv, 2008; Niv, 2009; Niv & Langdon, 2016). That RL appears to be so transparently embedded implies that it can be seen as a general cognitive mechanism and used in an immediate way to make hypotheses about and interpretations of a wealth of behavioral and neural data.

The availability of a growing suite of environments (from simulated robots to Atari games), toolkits, and sites for comparing and reproducing results about RL algorithms applied to a variety of tasks (Lazaridou, Peysakhovich, & Baroni, 2016; Havrylov & Titov, 2017; Evtimova, Drozdov, Kiela, & Cho, 2017; Jorge, Kågebäck, & Gustavsson, 2016) makes it possible to study cognitive science questions through a different lens. Cognitive science experiments are often carried out in real life settings involving questionnaires and surveys which are both costly and suffer from variability in responses. If RL algorithms are indeed a good proxy for actual human learning, then insights about questions of universals in language learning could be obtained very cheaply and reliably via controlled experiments in such *in silico* settings. We demonstrate this with a focus on questions about the universality of color categories and words in language.

This paper contributes a model of color partitioning based in recent advances in deep reinforcement learning that directly simulates an interactive, communication-based model of arriving at a color term consensus that partitions the color space for optimal, lowest-cost communication. We thus advance work in measuring the role of communicative cost by providing a procedure to generate color naming schemes that reflect those results. Our system is a starting point in developing models to test further hypotheses about universals in semantic partitioning while preserving a notion of interaction in the generation of meaning.

Evaluating color–word schemes

An assignment of a color term to each chip corresponds to a categorical partition of color space this arrangement represents. Given such a partition $\mathcal{P}$, Regier et al. (2007, 2015) propose quantitative schemes to test the hypothesis that attested color naming systems achieve a near-optimal trade–off between informativeness and complexity—that is, that they are nearly as informative as is theoretically possible for their level of complexity (i.e., for their number of color terms).

Partitions of color space

The first criterion, proposed in Regier et al. (2007), is a measure of color space partition quality. The objective function measures the extent to which such an assignment of category labels to chips maximizes similarity within categories and minimizes similarity across categories. In fact, this measure is well known in theoretical computer science and approximation algorithms as the (weighted) *correlation clustering* problem (Bansal, Blum, & Chawla, 2004; Giotis & Guruswami, 2006; Ailon, Charikar, & Newman, 2008). Given a graph $G = (V,E)$ with weights $w^+, w^- : E \to R$; the *maximizing agreements* version of the problem seeks to partition the vertices into disjoint sets so as to maximize

$$\sum_{cat(i)=cat(j)} w^+(i,j) + \sum_{cat(i)\neq cat(j)} w^-(i,j). \tag{1}$$

Here $cat(i)$ refers to the subset in the partition to which the vertex i belongs.

In our case, as in Regier et al. (2007), we take $w^+(i,j) := sim(i,j)$ and $w^-(i,j) := (1 - sim(i,j)$, where $sim(i,j)$, the similarity of two colors i and j, is adopted from the psychological literature on categorization:

$$sim(i,j) := \exp\left(-c dist(i,j)^2\right), \tag{2}$$

where $dist(x,y)$ is the CIELAB distance between colors x and y, and c is a scaling factor (set to 0.001 for all simulations reported here). It has a maximum value of 1 when chips x and y are the same (i.e., $dist(x,y) = 0$) and a value that falls off approaching 0 as the distance between chips x and y becomes arbitrarily large. This similarity function thus captures the qualitative observation that beyond a certain distance colors appear "completely different" so that increasing the distance has no further effect on dissimilarity.

Given this choice, the problem is to maximize over all partitions $\mathcal{P}$,

$$CC(\mathcal{P}) := \sum_{Cat(i)\neq Cat(j)} 1 - 2sim(i,j)$$

The problem is NP-hard but can be computed exactly for small sizes with an integer LP. Regier et al. (2007) propose a heuristic to approximate the optimum.

Information–theoretic analysis: communicating over a noisy channel

The second criterion proposed in Regier et al. (2015) takes an information–theoretic perspective via a scheme in the form of a speaker and a listener communicating over a noisy channel. The speaker attempts to communicate a color from the WCS color grid. Each agent maintains a belief in the form of a probability distribution over colors. The speaker's distribution s is concentrated on the single color i she is trying to convey by uttering a word w. After listening to the uttered word, the listener tries to reconstruct the color, resulting in a probability distribution ℓ over colors. The reconstruction error $e(i)$ for color i is the KL divergence $D(s||\ell)$ which equals the surprisal $-\log \ell(i)$. The listener distribution ℓ is concentrated on the category labeled with the word communicated and

$$\ell(t) = s(t) / \sum_{i \in Cat(t)} s(i)$$

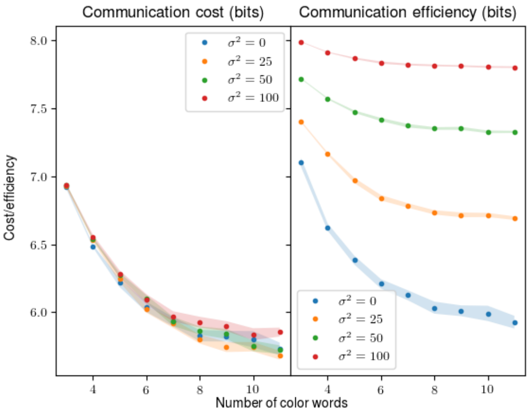

Figure 1: Communication cost (left) and communication efficiency (right) for different levels of noise as a function of the number of allowed color terms. The points indicate the mean values from 50 independent runs and the shaded area constitutes the std. deviation $\sigma/4$.

where

$$s(i) = \sum_{j \in Cat(i)} sim(i,j)$$

. This is motivated by an exemplar selection argument (i.e., from a category); one tends to select the most representative exemplar. To get an aggregate measure of the reconstruction error over all colors in the domain universe of colors, we define $n(i)$ as the need probability for target color i and compute the expected reconstruction error over all colors as

$$E := \sum_i n(i)e(i). \tag{3}$$

Hence, E measures the expected information loss incurred when transferring color information between two agents over a linguistic communication channel.

Communication efficiency

In the work of (Gibson et al., 2017) the related measure communication efficiency is defined as

$$\sum_c p(c) \sum_w p(w|c) \log_2 \frac{1}{p(c|w)} \tag{4}$$

and computes the expected average surprisal of each color chip c while assuming a uniform prior p(c).

Combined criterion

An alternative suggestion is that the listener selects uniformly from the category labeled by the received word. In this case, the surprisal is $\log_2 |Cat(w)|$. This suggests we optimize a combination of the two criteria, the quality of the color space and the surprisal, that is find a partition $\mathcal{P}$ to minimize

$$\begin{aligned} & \frac{1}{n}\sum_w |Cat(w)| \log_2 |Cat(w)| + CC(\mathcal{P}) \\ = \quad & \log n - \mathcal{H}(\mathcal{P}) + CC(\mathcal{P}). \end{aligned} \tag{5}$$

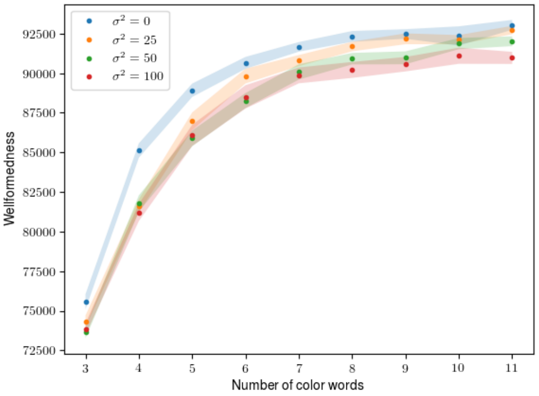

Figure 2: Wellformedness for different levels of noise, as a function of the number of words. The points indicate the mean values from 50 independent runs and the shaded area constitutes the std. deviation $\sigma/4$..

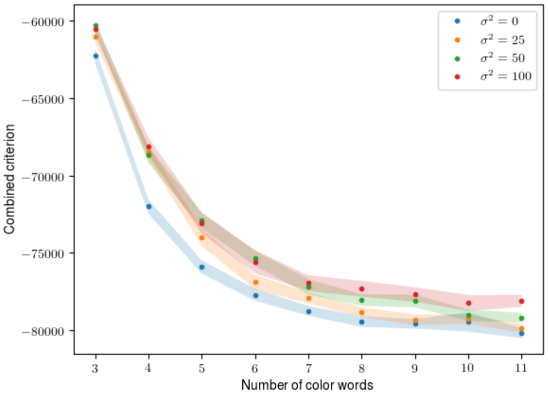

Figure 3: Combined criterion for different levels of noise as a function of the number of words.The points indicate the mean values from 50 independent runs and the shaded area constitutes the std. deviation $\sigma/4$.

where $\mathcal{H}(\mathcal{P})$ is the *entropy* of the partition $\mathcal{P}$ (with uniform measure on the elements).

A DeepColor framework for communication

We develop a version of the communication setup from Regier et al. (2015) via two automated agents implemented as feed forward neural networks trained via reinforcement learning. Though the structure of the model is more general, for the purposes of our game, the stimulus always consists of a color coordinate in CIELAB color space and the guess indicates which of the color chips in the palette the receiving agent believes the sender is trying to communicate. Both the sender and receiver is modeled using a *multilayer perceptron* with one hidden layer of $k = 20$ units.

The sender decides which message to send by sampling from its internal distribution over the set of color terms W

$$w \sim p(W|\mathbf{t};\Omega_s) = \mathrm{softmax}(\phi_s^T \tanh \theta_s^T \mathbf{t})$$

given the stimulus $\mathbf{t} = \mathrm{CIELAB}(c) + \beta$, where $c \sim U$ and the noise vector $\beta \sim N(0, \sigma^2)$. $\Omega_s = \{\theta_s \in \mathbb{R}^{k\times 3}, \phi_s \in \mathbb{R}^{|W|\times k}\}$ is the parameterization of the sender agent, and the softmax is the defined as $\mathrm{softmax}_j(\mathbf{z}) = e^{z_j}/\sum_i^{|\mathbf{z}|} e^{z_i}$. The bias terms have been omitted for brevity.

The receiver interprets the message and computes a distribution over all color tiles in U given the received message w as $p(U|\mathbf{w};\Omega_r) = \mathrm{softmax}(\phi_r^T \tanh \theta_r^T \mathbf{w})$, where $\Omega_r = \{\theta_r \in \mathbb{R}^{k\times|W|}, \phi_r \in \mathbb{R}^{|U|\times k}\}$ parameterize the receiver.

The sender part of the model is trained using the well known policy gradient method REINFORCE (Williams, 1992), an algorithm that aims to update the parameters of the model such that to maximize a given reward function. In our case this reward is chosen as the communication outcome $r := \mathrm{sim}(c, \mathrm{argmax}\, p(U|\mathbf{w};\Omega_r))$, where sim is the color similarity function defined in Equation 2. Plugging this reward into the REINFORCE cost function gives us

$$J_s(\Omega_s) = -\frac{1}{N_b}\sum_n^{N_b} \log p(W = w_n|\mathbf{t}_n;\Omega_s) * r_n.$$

where N_b corresponds to the number of games that the cost is computed over. The receiver objective can be modeled directly, i.e., without using a reward, as the surprisal incurred by the receiver after seeing the intended color chip c:

$$J_r(\Omega_r) = -\frac{1}{N_b}\sum_n^{N_b} \log p(U = c_n|\mathbf{w_n};\Omega_r).$$

The final objective function $J(\Omega_s, \Omega_r) = J_r(\Omega_r) + \lambda J_s(\Omega_s)$ combines the sender and receiver parts using the mixing factor $\lambda = 100$, and the model is trained by minimizing this objective. At the beginning of training all parameters $\{\Omega_s, \Omega_r\}$ are initialized to random values and the optimization $\min_{\Omega_s,\Omega_r} J(\Omega_s, \Omega_r)$ is performed using stochastic gradient decent with ADAM (Kingma & Ba, 2014). The batch size is set to $N_b = 100$ games, and the model is trained for a total of $N = 10000$ games.

After the model has been trained, the resulting color partitions $\mathcal{P}$ is extracted from the sender agent as its most probable message corresponding to each color tile $c \in U$.

Experiments

We conducted experiments to answer three primary questions: (1) Does DeepColor converge on high quality color partitions measured by both well–formedness of the resulting partitions and the information efficiency of the communication scheme; (2) how does varying noise during training influence the results in (1); and (3) does noise influence the number of words that the agents choose to employ?

To answer these questions we trained independent agents on all combinations of the number of allowed color terms $\in$

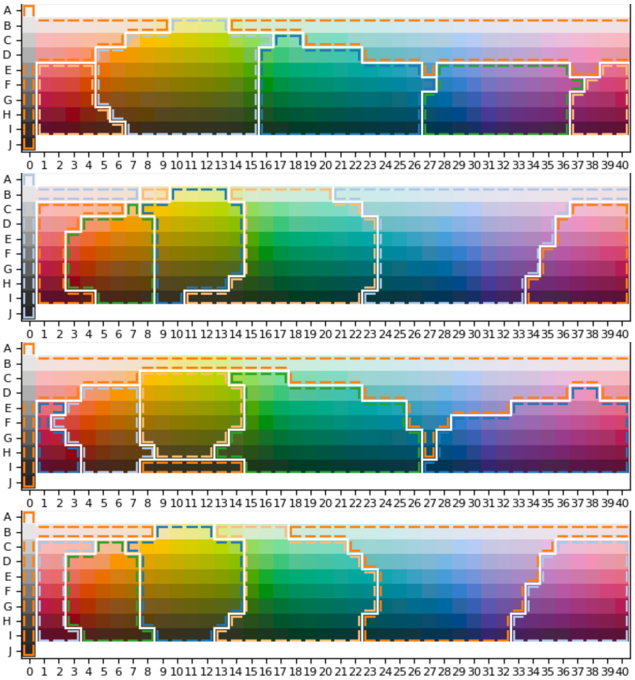

Figure 4: Resulting color maps when allowing the agents to use 5 color terms and adding different amounts of noise ($\sigma^2 \in \{0,25,50,100\}$) to the color chips. The color maps are presented in order of increasing noise from top to bottom.

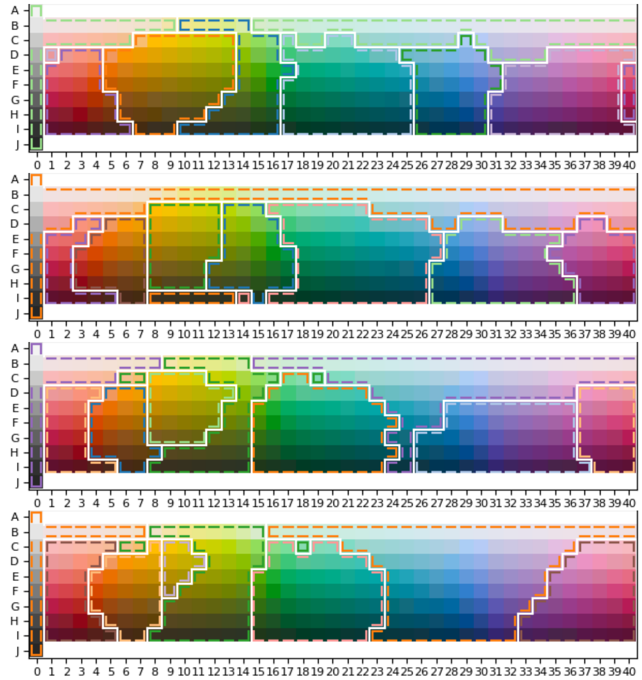

Figure 5: Resulting color maps when allowing the agents to use 11 color terms and adding different amounts of noise ($\sigma^2 \in \{0,25,50,100\}$) to the color chips. The color maps are presented in order of increasing noise from top to bottom. Note that the agents never choose to use all 11 color words but that the number they choose to use depend on the amount of noise added to the chips during training.

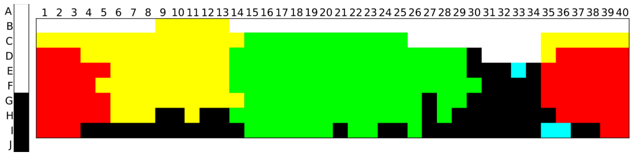

Figure 6: After Regier et al. (2015). WCS color map of the Iduna language, which has five color terms.

$\{3,4,\ldots,11\}$ and the variance of the noise added to the color chips $\sigma^2 \in \{0,25,50,100\}$.

In figures 1-3, we represent the performance of our models relative to the communication efficiency (see equation 4) communication cost (see equation 3), wellformedness (see equation 1), and combined criteria (see equation 5). The number of color words has minimal effect on efficiency, particularly at higher noise. As Gibson et al. (2017) suggest, this indicates the robustness of the WCS investigative paradigm: in our case, in an artificial agent setting. Our process is highly tolerant to noise, with the highest-noise scenario having similar overall cost reductions and well-formedness improvements relative to the number of words. More words represent a reduction in costs, with the curve reaching its asymptote at the higher number of words.

Figures 4 and 5 display sample color maps generated by the reinforcement learning process. A bounded region on each map represents a word that the system assigned to each color; they are not necessarily contiguous and sometimes wrap around the table. At 5 color terms, there is considerable stability among regions regardless of noise. At 11 color terms (this is a maximum, as the system is not obliged to assign all colors), we see more hierarchy and variation in the maps and more terms exploited as the noise increases. We also observe large similarities with the 5-color maps in terms of overall boundaries, with greater subdivisions.

Figure 6 presents a WCS color mode map of the Iduna language. Compared to the five-color maps of figure 4, we see large overlap in the color boundaries that appear. One major difference is that Iduna uses one color term for a small number of Munsell chips, in the manner that the noisiest two maps of figure 5 create singleton or small–group color terms.

Figure 7 shows that noise does have an effect on the number of words that the agents decided to employ—large amounts of noise decrease the active vocabulary of the agents, so precise terms may be less useful when referring to objects that vary in color. However, we also observe that no noise can have a similar detrimental effect which may be due to the noise acting as a stabilizing regularizer during training.

The presented example color maps have been sampled randomly from our experimental results that were compiled after training DeepColor using the hyper-parameters specified in the previous section.

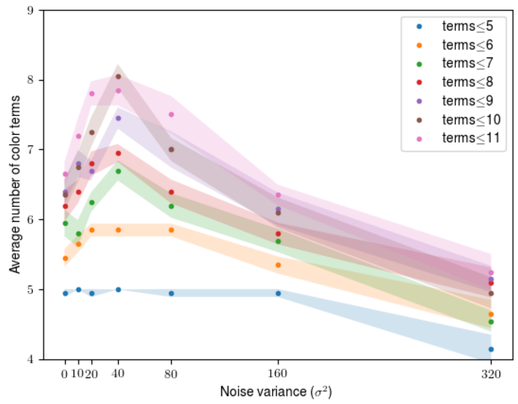

Figure 7: Average number of words actually used by the agents after training under different amounts of noise. The points indicate the mean values from 20 independent runs and the shaded area constitutes the std. deviation $\sigma/4$.

Conclusions and future work

From where do apparent regularities in color naming patterns across language arise? A recent strain of research explains this via a perceptual well-formedness criterion as well as communicative cost. Implicit in communicative cost is the role of linguistic interaction. Our deep reinforcement model directly implements an interactional learning model, and the color term maps generated by this model replicate empirical observations in human–language color mapping.

Compared to other semantic domains, color is a simple test bed for interaction–based approaches to modeling the partitioning of meaning spaces. The initial success of this model not only allows for further experiments and comparisons in the quality and grounding of color partitions, such as through image data, but application to other semantic domains, such as the partitioning of WordNet synsets. It also demonstrates the viability of reinforcement learning in implementing information–theoretic approaches to representing semantic distinctions in linguistic interaction.

Acknowledgments The authors would like to acknowledge the project *Towards a knowledge-based culturomics* supported by a framework grant from the Swedish Research Council (2012–2016; dnr 2012-5738) as well as the SRC-funded Centre for Linguistic Theory and Studies in Probability (CLASP) at the University of Gothenburg.

References

Ailon, N., Charikar, M., & Newman, A. (2008). Aggregating inconsistent information: Ranking and clustering. *J. ACM*, *55*(5), 23:1–23:27.

Bansal, N., Blum, A., & Chawla, S. (2004). Correlation clustering. *Machine Learning*, *56*(1-3), 89-113.

Berlin, B., & Kay, P. (1969). *Basic color terms: Their university and evolution*. California UP.

Dayan, P., & Niv, Y. (2008). Reinforcement learning: The good, the bad and the ugly. *Current Opinion in Neurobiology*, *18*(3), 1-12.

Evtimova, K., Drozdov, A., Kiela, D., & Cho, K. (2017). Emergent language in a multi-modal, multi-step referential game. *CoRR*, *abs/1705.10369*.

Gibson, E., Futrell, R., Jara-Ettinger, J., Mahowald, K., Bergen, L., Ratnasingam, S., ... Conway, B. R. (2017). Color naming across languages reflects color use. *Proc Natl Acad Sci USA*, *114*(40), 1078510790.

Giotis, I., & Guruswami, V. (2006). Correlation clustering with a fixed number of clusters. *Theory of Computing*, *2*.

Havrylov, S., & Titov, I. (2017). Emergence of language with multi-agent games: Learning to communicate with sequences of symbols. In I. Guyon et al. (Eds.), *Advances in neural information processing systems 30* (pp. 2146–2156). Curran Associates, Inc.

Jorge, E., Kågebäck, M., & Gustavsson, E. (2016). Learning to play guess who? and inventing a grounded language as a consequence. *CoRR*, *abs/1611.03218*.

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Lazaridou, A., Peysakhovich, A., & Baroni, M. (2016). Multi-agent cooperation and the emergence of (natural) language. *CoRR*, *abs/1612.07182*.

Lucy, J. A. (1997). The linguistics of color. In C. L. Hardin & L. Maffi (Eds.), *Color categories in thought and language* (pp. 320–346). Cambridge University Press.

Niv, Y. (2009). Reinforcement learning in the brain. *The Journal of Mathematical Psychology*, *53*(3), 139-154.

Niv, Y., & Langdon, A. (2016). Reinforcement learning with marr. *Current Opinion in Behavioral Sciences*, *11*(3).

Piantadosi, S. T., Tily, H., & Gibson, E. (2011). Word lengths are optimized for efficient communication. *Proceedings of the National Academy of Sciences*, *108*(9), 3526–3529.

Regier, T., Kay, P., & Khetrapal, N. (2007). Color naming reflects optimal partitions of color space. *Proc Natl Acad Sci USA*, *104*(3), 1436-1441.

Regier, T., Kemp, C., & Kay, P. (2015). Word meanings across languages support efficient communication. In B. MacWhinney & W. O'Grady (Eds.), *The handbook of language emergence* (p. 237-263). Hoboken NJ: Wiley-Blackwell.

Saunders, B. (1995). Disinterring basic color terms : a study in the mystique of cognitivism. *History of the Human Sciences*, *8*(4), 19-38.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning an introduction*. MIT Press.

Wiering, M., & van Otterlo, M. (Eds.). (2012). *Reinforcement learning: State-of-the-art*. Springer.

Williams, R. J. (1992). Simple statistical gradient-following algorithms for connectionist reinforcement learning. In *Reinforcement learning* (pp. 5–32). Springer.