Adaptive value within natural language discourse*

Michael L. Best
Georgia Institute of Technology

A trait is of adaptive value if it confers a fitness advantage to its possessor. Thus adaptiveness is an ahistorical identification of a trait affording some selective advantage to an agent within some particular environment. In results reported here we identify a trait within natural language discourse as having adaptive value by computing a trait/fitness covariance; the possession of the trait correlates with the replication success of the trait's possessor. We show that the trait covaries with fitness across multiple unrelated discursive groups. In our analysis the trait in question is a particular statistically derived word-in-context, that is, a word set. Variation of the word-usage is measured as the relative presence of the word set within a particular text, that is, the percentage of the text devoted to this set of words. Fitness is measured as the rate in which the text is responded to, or replicates, within an online environment. Thus we are studying the micro-evolutionary dynamics of natural language discourse.

Keywords: evolution in communication, adaptation, population memetics, cultural evolution

Introduction

Computational studies of the evolution of communication have concentrated on three main areas.

- The first is the evolution of signals and development of shared lexicons. Researchers have constructed systems in which agents evolve cooperative signaling strategies and emergent shared lexicons (e.g. Werner & Dyer 1992; MacLennan 1991; Ackley & Littman 1994; Steels 1996; Saunders & Pollack 1996).

Interaction Studies 7:1 (2006), **1–15**.
ISSN 1572–0373 / E-ISSN 1572–0381 © John Benjamins Publishing Company

- Second, researchers have examined the evolution of phonology and syntax, for instance the self-organization of vowel systems (de Boer 1997) and word clusters (Hashimoto 1994).
- Finally, a few studies have used computational methods to examine innateness and the critical period of language learning (Hurford 1991; Batali 1994).

We are researching, in contrast, computational models of the adaptive value of word usage within natural language discourse. Thus, we are engaging in a micro-evolutionary analysis of the dynamics of natural language. In contrast to work considering the development of communication itself (in particular morphosyntactic), the establishment of shared lexicons, or the acquisition of language, we are examining the evolutionary dynamics of established language use over short, proximal, time periods.

While we know of no other micro-evolutionary consideration of natural language, such phenomena within culture at large have been studied via both computational and formal modeling by a number of researchers. In a nice turn of phrase, Dawkins (1976) coined the term "meme" to describe a particulate cultural replicator and a collection of researchers have studied such cultural replicators through simulation environments (e.g. Gabora 1995; Best 1999). A number of substantial theories of cultural transmission have been developed (Lumsden & Wilson 1981; Cavalli-Sforza & Feldman 198l; Boyd & Richerson 1985; Barkow 1989; Durham 1991; reviewed in Durham 1990). Many of these models do develop a micro-evolutionary theory of the transmission of particulate cultural replicators between populations of social agents. But none of these works have studied human natural language directly, and only rarely do they study cultural phenomena via empirical analysis.

In results reported here we establish that a trait, a word set used in a particular way, is of adaptive value. That is, we show that the relative presence of the particular word-usage covaries with the rate in which its body of text is replicated within an online text environment. To some this ahistorical demonstration of fitness enhancement is enough to label the trait an adaptation (Clutton-Brock & Harvey 1979; Reeve & Sherman 1993; but compare Gould 1984). For others (Waddington 1957; Williams 1966; Lewontin 1978; Sober 1984; but compare Bock 1980) a history-laden investigation of the trait is critical. In particular, the trait must have established itself in trans-generational time due to some design quality it possesses relative to variant forms.

This work describes a model and computational framework with which to study micro-evolutionary dynamics within discourse. We believe this opens up a broad area for potential continued study and progress.

The natural language corpora

The area of culture we study is natural language discourse which we gather from discussions posted to the USENET News (NetNews) system. NetNews is a popular computer-based discussion system supported through standardized protocols on the Internet (Kantor & Lapsley 1986). Articles are authored by interested users and posted to particular discussion groups called newsgroups. Newsgroups are organized in a tree hierarchy which has at its root a top-level category and moves to more specific topics as you progress towards the leaves (see Figure 1). As an example the "sci.bio.evolution" newsgroup concerns itself with questions of evolution as a refining category of biology ("bio"), which in turn is a refining category of the set of scientific discussion groups ("sci"). Posts to NetNews can be independent messages, but quite frequently are follow-up messages to previous posts. The software system threads together these follow-up messages and can automatically include the previous message's text within the body of the follow-up post.

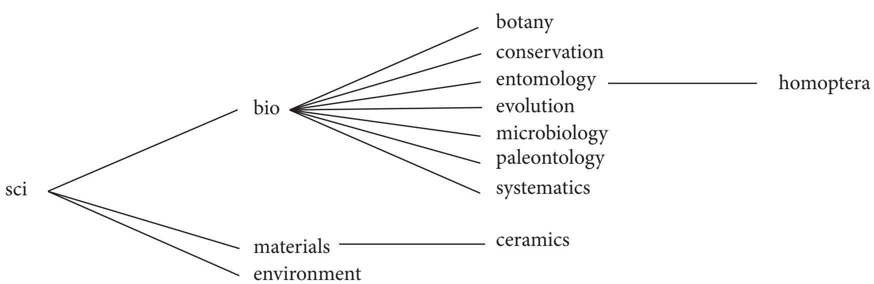

Figure 1. A portion of the NetNews newsgroup hierarchy. Moving from left to right we have newsgroups like sci.bio and sci.bio.botany.

Posts are composed of a number of fields, only a few of which are relevant here. The user creating the post is responsible for the post body, that is, the actual text of the message, as well as a subject line. The subject line is composed of a few words which describe what the post is about. NetNews software will append to posted messages a number of additional fields including a timestamp and the user name of the person who created the post.

```
From: mikeb@gatech.edu (Michael Best)
Newsgroups: sci.bio.evolution
Subject: Language use of Homo habilis
Date: 26 Feb 1998 02:17:05 -0700
Can anyone point me to studies on the
likelihood of language use in Homo habilis?
```

Texts from the NetNews system have a number of favorable properties which lend themselves to the analysis of micro-evolutionary phenomena. First, each post is automatically timestamped such that a clear and sufficiently accurate arrow of time exists through the texts. Next, the follow-up message facility describes lineages within the corpora. A post which receives a follow-up will, with high probability, pass on a significant set of traits to its follow-up message in the form of repeating word co-occurrences. This occurs both due to the system's automatic inclusion of the previous message's text as well as through the action of the human authors. The result is that a follow-up message is highly likely to share considerable traits with its parent message relative to the population of messages as a whole; in other words this produces a lineage with heredity (Dawkins 1982).

Units of selection

To demonstrate adaptive value we must first identify traits which are replicating though the corpus (Dawkins 1982; Hull 1988). Thus, we must seek elements within the corpus which are:

- repeating,
- with sufficient copying fidelity,
- but not with perfect fidelity, there must be some variation (Lewontin 1970; Eigen 1973; Dawkins 1982).

Elsewhere (Pocklington & Best 1997; Best 1997; Best 1998a) we have argued that a statistical text-retrieval technique, based on the vector space representation and Latent Semantic Indexing (Deerwester, Dumais, Furnas, Landauer & Harshman 1990), satisfies these desiderata. That is, we argue that statistically derived word sets are traits within discourse.

Vector space representation

We will now overview this statistical text-retrieval technique. We begin with a corpus composed of the full-text of a group of posts. We analyze the corpus and identify a high-dimensioned space which describes the conceptual elements within the texts. For each post we identify a point within this space which captures it semantically. This technique is known as a vector space representation (Salton & McGill 1983). Each dimension in this space will represent a term from the corpus where a term is a word that occurs with some frequency (e.g. in at least three posts) but not with too much frequency (e.g. the word "not" is dropped from the term list). The goal is to arrive at a set of terms which semantically capture the texts within the corpus.

Given the conceptual space described by this set of terms each post can be represented as a point within this space. We score each document according to the frequency of each term within its text, and assign each term/document pairing this *term weight*. The weighting we use for each term/document pair is a function of the *term frequency* (simply the number of times the term occurs in the post) and the *inverse document frequency (IDF).* Consider a corpus of m posts and a particular term, j, within a list of n terms. Then the IDF is given by,

$$IDF_j = \log\left(\left\lfloor \frac{m - m_j}{m_j} \right\rfloor\right),$$

where m_j is the number of posts across the entire corpus in which term j appears. The term weight for a document, i, and term, j, is then defined by,

$$TermWeight_{ij} = w_{ij} = \log(TermFrequency_{ij}) \cdot IDF_j.$$

Thus, each term weight is a function of the inter- and intra-document term frequencies.

Each post, i, is now represented by a particular term vector,

$$r_i = (w_{i1}, w_{i2}, \ldots, w_{in}).$$

The entire collection of m term vectors, one for each post, define the term/document matrix, A,

$$A = \begin{bmatrix} r_1 \\ r_2 \\ \cdots \\ r_m \end{bmatrix} = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ w_{m1} & w_{m2} & \cdots & w_{mn} \end{bmatrix}.$$

This set of steps, culminating in the term/document matrix, form the basis for much of modern text retrieval or filtering and are at the core of most Web search engines. (See Frakes & Baeza-Yates (1992) for a review of modern text retrieval techniques.)

Latent semantic indexing

LSI is a technique to distill high-order structures from a term/document matrix. This structure consists of sets of terms which re-occur together through the corpus with appreciable frequency. The re-occurring term sets are discovered through a principal component method called Singular Value Decomposition (SVD). While LSI was primarily developed to improve text retrieval, we are interested in its ability to find replicating term sets. These describe re-occurring words-in-context and thus identify evolutionarily significant traits. We will first overview the LSI technique and then discuss how it discovers replicators.

LSI was originally proposed and has been extensively studied by Susan Dumais and her colleagues (Furnas, *et al.* 1988; Deerwester, *et al.* 1990; Dumais 1992, 1993). Peter Foltz investigated the use of LSI in clustering NetNews articles for information filtering (Foltz 1990). Michael Berry and co-authors researched a variety of numerical approaches to efficiently perform SVD on large sparse matrices such as those found in text retrieval (Berry 1992; Berry, *et al.* 1993; Berry & Fierro 1995).

The SVD technique decomposes the term/document matrix into a left and right orthonormal matrix of eigenvectors and a diagonal matrix of eigenvalues. The decomposition is formulized as,

$$A \approx A_k = U \Sigma V^T = \sum_{i=1}^{k} u_i \cdot \sigma_i \cdot v_i^T .$$

Here, we see that the term/document matrix, A, is approximated by a rank-k decomposition, A_k; in fact the SVD technique is known to produce the best rank-k approximation to a low-rank matrix (Berry 1992).

We are interested in only the right orthonormal matrix of eigenvectors, V^T. Each row of this matrix defines a set of terms whose co-occurrence have some statistically salient but not perfect re-occurrence throughout the corpus. That is, each eigenvector describes a subspace of the term vector space for which the terms are frequently found together. These *term-subspaces* describe a set of semantically significant associative patterns in the words of the underlying corpus of documents; we can think of each subspace as a *conceptual index* into the corpus (Furnas *et al.* 1988).

For instance, in an analysis of military posts we might find that three words, "harbor", "japan", and "pearl" re-occur together with statistical significance. Therefore these words-in-context, or particular semantic trait, are replicating with success. It is these term-subspaces which describe our replicating traits and allow for an evolutionary analysis because they meet our desideratum. Elsewhere we have argued that these co-occurring word sets are an evolutionary unit of selection (Pocklington & Best 1997; Best 1998a). We have also studied various ecological phenomena within our text corpora, for instance competition and mutualism (Best 1997).

Our final text analysis step is to "compress" the original term/document matrix by multiplying it with this right orthonormal matrix of eigenvectors (in other words we perform a projection). This, in effect, produces a *term-subspace/document* matrix. Each post is represented by a collection of weights where each weight now describes the degree to which a term-subspace is expressed within its post's text.

Adaptive value

If a vector of term-subspace weights describe a set of traits for each text then, with the addition of a measure of fitness, we will be able to compute each trait's adaptive value as a trait/fitness covariance. We consider the fitness of a post lineage to be the number of offspring for that lineage as a function of time (Crow & Kimura 1970). Thus, the more replies within a thread for some unit of time the higher the relative fitness of that lineage. And we wish to determine if that relative fitness covaries with the degree to which the trait is expressed, that is with the varying expression of the trait.

Consider a particular lineage of texts along with some particular term-subspace metric trait. Then this trait taken across our population of texts describes a time series; each time a new text is posted to the thread we can plot against time its weight for the particular trait. Moreover, our measure of fitness also describes a time series; we can plot against time the current fitness (number of posts) for each block of time. (Figures 2–4 are just such plots.) Our mechanism for computing adaptive value should now be evident; we look for instances where the trait/fitness correlation coefficient, in other words the covariation between these two time-series, is relatively high.

Empirical analysis

We have identified the trait/fitness correlations for a particular trait within three different sets of texts. These corpora were composed of all posts to three different newsgroups over a set of days. The three newsgroups were sci.skeptic (scientific issues, skeptical attitude), soc.subculture.bondage-bdsm (sex,

Table 1. Three corpora gathered from differing newsgroups. The "Nazi" trait shows high correlation with fitness while the average correlation for all traits is near zero.

Newsgroups	sci.skeptic	soc.subculture. bondage-bdsm	alt.politics.usa. constitution
Total number of posts to newsgroup	11,758	1,160	494
Number of posts in particular lineage (thread)	101	66	29
Dates	9/20/95–9/26/95	9/28/97–10/6/97	10/30/97–11/2/97
"Nazi" trait/fitness correlation coefficient	0.8408 ($p < 0.001$)	0.8166 ($p < 0.001$)	0.5308 ($p < 0.001$)
Average trait/fitness correlation coefficient	–0.0102	0.0147	0.0112

bondage, and discipline), and alt.politics.usa.constitution (discussions related to the US constitution). The size of the corpora ranged from 494 to 11,758 texts gathered over the course of 4 to 9 days (see Table 1). In previous work (Pocklington & Best 1997; Best 1998) we had discovered a very successful trait within the sci.skeptic corpus, to wit a highly successful replicating word set contained the three terms, "John", "Smith", and "Nazi". A very large thread within this corpus centered around a particular person, John Smith (we use a pseudonym here), and consisted of a heated discussion of Smith's posting habits. A very frequent comment (thus the high replication rate) was that Smith was a "Nazi". In our original study we found this Nazi trait (namely the word set "John, Smith, Nazi") to have a very high trait/fitness correlation coefficient (Figure 2 and Table 1). Thus, the more often a text referred to John Smith as a Nazi the more follow-up messages the post would receive (which, in turn, would lead to more copies of the trait since those follow-ups invariably would call Smith a Nazi one way or another).

We were interested in finding other similar uses of the term Nazi; that is, uses of the word as name-calling and not with reference to German National-Socialists[1]. We easily found two other instances of a "Nazi" trait which enjoyed high trait/fitness correlations. In both cases a term-subspace was discovered in which people were called Nazis as a pejorative without reference to National-Socialists (this was verified by inspection). In Table 1 we list the trait/fitness correlation coefficient for the "Nazi" trait within each corpus as well as the mean trait/fitness correlation for every other replicator found in the texts. These results show three different lineages in three different corpora (from politics to bondage to skepticism) which all employ the same trait (calling people Nazis)

sci.skeptic

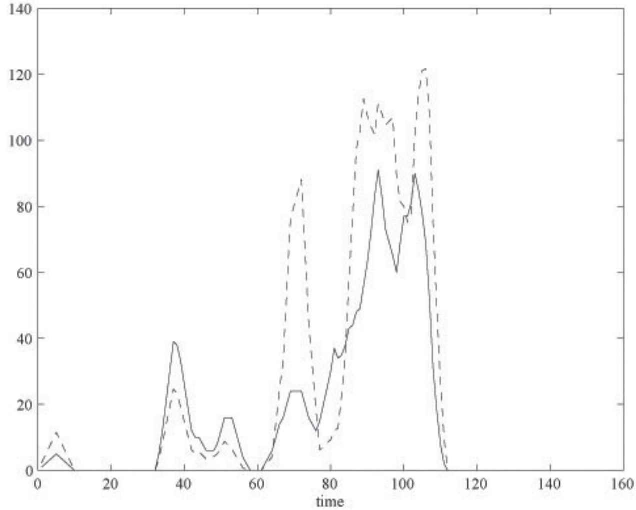

Figure 2. Trait (solid line) and fitness (dashed line) is shown to co-vary with time for posts to sci.skeptic. The particular trait shown is the "Nazi" replicator. The Y-axis is scaled relative fitness and metric trait weight. Fitness and trait weight drop to zero at rightmost of graph as an artifact of the corpus being exhausted.

soc.subculture.bondage-bdsm

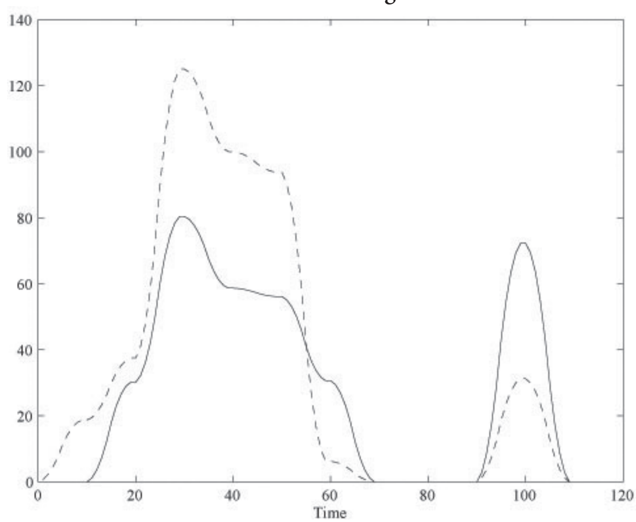

Figure 3. Trait (solid line) and fitness (dashed line) is shown to co-vary with time for posts to soc.subculture.bondage-bdsm. The particular trait shown is the "Nazi" replicator. The Y-axis is scaled relative fitness and metric trait weight. Fitness and trait weight drop to zero at rightmost of graph as an artifact of the corpus being exhausted.

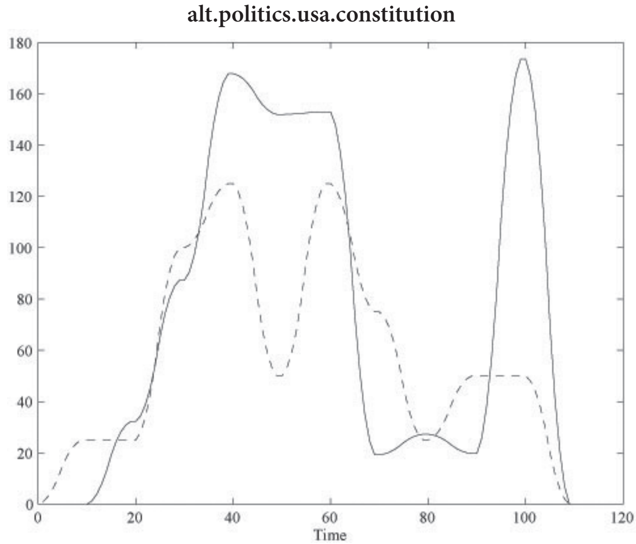

Figure 4. Trait (solid line) and fitness (dashed line) is shown to co-vary with time for posts to alt.politics.usa.constitution. The particular trait shown is the "Nazi" replicator. The Y-axis is scaled relative fitness and metric trait weight. Fitness and trait weight drop to zero at rightmost of graph as an artifact of the corpus being exhausted.

with considerable adaptive success (Figure 2–4 and Table 1). Note that we did not locate a large number of occurrences of this trait across many newsgroups and then systematically select those occurrences with high adaptive value. We selected the first two instances of the trait we identified in our exploration of posts and both of them had a significant trait/fitness covariance; in other words, we did not simply select the tail of some distribution of covariances.

To summarize, a particular word-usage, "Nazi" as a general perjorative attack word, is identified in three sets of texts. This word set is a trait of these texts, and is found through the SVD to replicate with statistical salience. The variants of this trait are the real-valued metric term weights. A value of zero means the text does not make use of the trait; as the value approaches one a larger proportion of the text is devoted to the trait. We find that as this metric value for the trait increases, so does the text's fitness — a trait/fitness covariance. Simply put, the more a text used "Nazi" to attack someone, the more responses the text would enjoy which in turn replicated the trait. Identifying this covariance demonstrates the trait's adaptive value (indeed Reeve & Sherman (1993) would argue it demonstrates an adaptation). Identifying the same adaptivness across multiple groups responding to similar selective pressures builds evidence that this trait is a true adaptation (Lewontin 1978).

Conclusions

We have developed a model for studying micro-evolutionary dynamics and population memetics within natural language discourse. Through statistical methods, based on principal component analysis, we find replicating sets of words within a corpus of texts posted to NetNews. That is, we recognize structures within the corpus which "segregate and recombine with appreciable frequency" (Williams 1966:24). Each text within the NetNews system acts in the role of vehicle for its collection of replicating memes (Dawkins 1982). For each of these posts we derive a vector representation where each dimension of the vector specifies a particular term-subspace, in other words one of the replicating word sets. Given a particular text and a particular term-subspace a real number is computed which measures the salience of that word set in that text. Thus, the vector space representation describes a string of metric traits (a memotype if you will) for each text. The newsgroups define an environment which the texts occupy and the human authors are a scarce resource and contribute to the cultural selective environment. We claim that this model provides an evolutionary theoretic ecological description of interacting texts within natural language discourse.

In results reported here, we have applied this model to an empirical consideration of adaptive value within text. We have found that the use of "Nazi" as a name-calling device has high adaptive value since we have found it to have a strong trait/fitness covariance within multiple groups of texts. But to claim that this use of "Nazi" is an adaptation requires, for most researchers, linking the trait to its history (Waddington 1957; Williams 1966; Lewontin 1978; Sober 1984). Have there been variants in usage of "Nazi" from which this particular pejorative name-calling usage has been selected and developed over generations? To answer this question we must add to our micro-evolutionary analysis a macro-evolutionary consideration of special design through semantic change.

In work currently in progress we are studying the semantic change of "Nazi". We note that while computational linguistics has made great strides in areas such as parsing, text analysis, automatic translation, and comparative and genetic analysis, there has been little application of computational linguistics towards problems of semantic change. Part of this is, no doubt, due the paucity of data. Online corpora such as those accumulated by Netnews are a relatively new phenomena. And without historical online corpora (here on the order of fifty years) it is difficult to engage in the sort of computational analysis of historical semantic change we require.

To conclude, the occurrence of "Nazi" as a pejorative attack name has strong adaptive value as demonstrated by its large trait/fitness covariance among multiple groups within NetNews. We demonstrated through computational micro-evolutionary analysis that as the particular word-usage becomes more salient, the replication rate for the text increases. This was shown to be true in three different collections of texts dealing with three fairly different sets of subject areas.

Notes

* We thank Warren Sack and Brian Smith for their commenting on a draft of this paper. Richard Pocklington has made significant contributions to the ideas in this paper.

1. We note our personal disapproval of the use of the word "Nazi" as a general pejorative as exhibited within these texts.

References

Ackley, D.H. & Littman, M.L. (1994). Altruism in the evolution of communication. In R.A. Brooks & P. Maes (Eds), *Artificial Life IV* (pp. 40–48). Cambridge, MA: MIT Press.

Barkow, J.G. (1989). *Darwin, sex, and status: Biological approaches to mind and culture*. Toronto: University of Toronto Press.

Batali, J. (1994). Innate biases and critical periods: Combining evolution and learning in the acquisition of syntax. In R. Brooks & P. Maes (Eds.), *Artificial Life IV* (pp. 160–171). Cambridge, MA: MIT Press.

Berry, M.W. (1992). Large-scale sparse singular value computations. *Journal of Supercomputer Applications* 6: 13–49.

Berry, M., Do, T., O'Brien, G., Krishna, V. & Varadhan, S. (1993). *SVDPACKC (Version 1.0) user's guide*. University of Tennessee Computer Science Department Technical Report, CS–93–194.

Berry, M.W. & Fierro, R.D. (1995). *Low-rank orthogonal decompositions for information retrieval applications*. University of Tennessee Computer Science Department Technical Report, CS–95–284.

Best, M.L. (1997). Models for interacting populations of memes: competition and niche behavior. *Journal of Memetics* [Online] 2. Available: http://www.cpm.mmu.ac.uk/jom-emit.

Best, M.L. (1998a). An ecology of text: using text retrieval to study alife on the net. *Journal of Artificial Life,* 3(4): 261–287.

Best, M.L. (1999). How culture can guide evolution: An inquiry into gene/meme enhancement and opposition. *Journal of Adaptive Behavior, 7*, 281–298.

Bock, W.J. (1980). The definition and recognition of biological adaptation. *American Zoologist, 20*, 217–227.

Boyd, R. & Richerson, P.H. (1985). *Culture and the evolutionary process.* Chicago: University of Chicago Press.

Cavalli-Sforza, L. & Feldman, M. (1981). *Cultural transmission and evolution: A quantitative approach.* Princeton, NJ: Princeton University Press.

Clutton-Brock, T.H., & Harvey, P.H. (1979). Comparison and adaptation. *Proceedings of the Royal Society of London B, 205*, 547–565.

Crow, F. & Kimura, M. (1970). *An introduction to population genetics theory.* New York: Harper and Row.

Dawkins, R. (1976). *The selfish gene*. Oxford: Oxford University Press.

Dawkins, R. (1982). *The extended phenotype.* San Francisco: W.H. Freeman.

de Boer, B. (1997). Generating vowel systems in a population of agents. In P. Husbands & I. Harvey (Eds.), *Fourth European Conference on Artificial Life* (pp. 503–510), Cambridge, MA: MIT Press.

Deerwester, S. Dumais, S.C., Furnas, G.W., Landauer, T.K. & Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American Society for Information Science, 41*, 391–407.

Dumais, S.T. (1992). LSI meets TREC: A status report. In *The First Text REtrieval Conference (TREC–1),* D. Harman (Ed.). NIST Special Publication 500–207.

Dumais, S.T. (1993). Latent semantic indexing (LSI) and TREC–2. In *The Second Text REtrieval Conference (TREC–2),* D. Harman (Ed). NIST Special Publication 500–215.

Durham, W.H. (1990). Advances in evolutionary culture theory. *Annual Review of Anthropology*, *19*, 187–210.

Durham, W.H. (1991). *Coevolution: Genes culture and human diversity.* Stanford, CA: Stanford University Press.

Eigen, M. (1973. *Steps towards life: An evolutionary approach.* Reading, MA: Addison-Wesley, Reading MA.

Foltz, P.W. (1990). Using latent semantic indexing for information filtering. *Proceedings of the 5th Conference on Office Information Systems*. ACM SIGOIS Bulletin 11(2–3).

Frakes, W.B. & Baeza-Yates, R. (Eds) (1992). *Information retrieval: Data structures & algorithms*. Upper Saddle River, NJ: Prentice Hall.

Furnas, G.W., Deerwester, S., Dumais, S.T., Landauer, T.K., Harshman, R.A., Streeter, L.A. & Lochbaum, K.E. (1988). Information retrieval using a Singular Value Decomposition model of latent semantic structure. *Proceedings of the 11th International Conference on Research and Development in Information Retrieval (SIGIR).* New York: Association for Computing Machinery.

Gabora, L. M. (1995). Meme and variations: A computational model of cultural evolution. In D. Stein (Ed) *1993 Lectures in complex systems.* Redwood City, CA: Addison-Wessley.

Gould, S.J. (1984). Covariance sets and ordered geographic variation in *Cerion* from Aruba, Bonaire and Curacao: A way of studying nonadaptation. *Systematic Zoology, 33*, 217–237.

Hashimoto, T. (1994). Usage-based structuralization of relationships between words. In P. Husbands & I. Harvey (Eds.), *Proceedings of the Fourth European Conference on Artificial Life, ECAL97* (pp. 483–492). Cambridge, MA: MIT Press/Bradford Books.

Hull, D.L. (1988). *Science as a process.* Chicago: University of Chicago Press.

Hurford, J.R. (1991). An approach to the phylogeny of the language faculty. In J.A. Hawkins & M. Gell-Mann (Eds.), *The evolution of human languages* (pp. 273–303). Redwood City, CA: Addison-Wesley Publishing Company.

Kantor, B. & Lapsley, P. (1986). *Network news transfer protocol: A proposed standard for the stream-based transmission of news.* Internet RFC–977.

Lewontin, R.C. (1970). The units of selection. *Annual Review of Ecology and Systematics, 1*, 1–18.

Lewontin, R.C. (1978). Adaptation. *Scientific American, 239*, 212–230.

Lumsden, C.J. & Wilson, E.O. (1981). *Genes, mind and culture: The coevolutionary process.* Cambridge, MA: Harvard University Press.

MacLennan, B. (1991). Synthetic ethology: An approach to the study of communication. In C.G. Langton, C. Taylor, H.D. Farmer, & S. Rasmussen (Eds). *Artificial life II: Santa Fe Institute studies in the sciences of complexity, Volume X*. Redwood City, CA: Addison-Wesley.

Pocklington, R. & Best, M.L. (1997). Cultural evolution and units of selection in replicating text. *Journal of Theoretical. Biology, 188*, 79–87.

Reeve, H.K & Sherman, P.W. (1993). Adaptation and the goals of evolutionary research. *Quarterly Review of Biology, 68*, 1–32.

Salton, G. & McGill, M.J. (1983). *Introduction to modern information retrieval.* New York: McGraw-Hill.

Saunders, G.M. & Pollack, J.B. (1996). The evolution of communication schemes over continuous channels. In P. Maes, M.J. Mataric, J. Meyer, J. Pollack, & S. Wilson (Eds.), *From animals to animats 4: Proceedings of the fourth international conference on simulation of adaptive behavior* (pp. 580–589). Cambridge, MA: MIT Press.

Sober, E. (1984). *The nature of selection: Evolutionary theory in philosophical focus.* Cambridge, MA: MIT Press.

Steels, L. (1996). Emergent adaptive lexicons. In P. Maes, M.J. Mataric, J. Meyer, J. Pollack, & S. Wilson (Eds), *From animals to animats 4: Proceedings of the fourth international conference on simulation of adaptive behavior* (pp. 562–567). Cambridge, MA: MIT Press.

Waddington, C.H. (1957). Selection by, of, and for. In C.H. Waddington (Ed.). *The Strategy of the genes* (pp 59–108). London: Allen and Unwin.

Werner, G.M. & Dyer, M.G. (1991). Evolution of communication in artificial organisms. In C.G. Langton, C. Taylor, H.D. Farmer, & S. Rasmussen (Eds). *Artificial life II: Santa Fe Institute studies in the sciences of complexity, volume X*. Redwood City, CA: Addison-Wesley.

Williams, G.C. (1966). *Adaptation and natural selection*. Princeton, NJ: Princeton Univ. Press.

Author's address:

Michael L. Best
Georgia Institute of Technology
781 Marietta St.
Atlanta, GA 30332
USA

mikeb@cc.gatech.edu

About the author

Michael L. Best is Assistant Professor with the Sam Nunn School of International Affairs at Georgia Tech and Adjunct Assistant Professor with their College of Computing. In addition he is a Fellow of the Berkman Center for Internet and Society at Harvard University and Research Affiliate with the Center for Technology, Policy, and Industrial Development and the Program for Internet & Telecoms Convergence at MIT. Michael is co-founder and Editor-in-Chief of *Information Technologies and International Development* published by the MIT Press. Michael serves as a frequent consultant to the World Bank, ITU, and USAID. He holds a Ph.D. from the MIT Media Lab and has served as Director of Media Lab Asia in India and head of the eDevelopment group at the MIT Media Lab.