CAMBRIDGE

Cambridge University Press & Assessment
978-1-107-00429-0 — Distributional Semantics
Alessandro Lenci , Magnus Sahlgren
Index
More Information

Index

abstract words, 272
across-the-board generalization, 350–351
affective content, 251–253
  Affective Distributional Hypothesis, 252
  centroid vector model, 252
  emotion, 251
  semantic orientation score, 252
  sentiment embedding, 253
  valence, 251
algebraic hypothesis of human mind, 349–350
amodal representation, 217
analogical generalization, 350–351
analogy, 182, 185
  3CosAdd, 184
  3CosMult, 184
  offset method, 183–185
  parallelogram model, *see* offset method
anisotropy, 336–337
antonymy detection, 240–243
  antonymy-specific embedding, 242
  APAnt, 241
  ATTRACT-REPEL, 242
  counterfitting, 241
  PairClass, 241
  pattern-based approach, 240
    Co-occurrence Hypothesis, 240
  Polarity Inducing LSA, 242
  SGNS with lexical contrast objective, 242
  vector-based approach, 241–243
artificial neural network, 21, 96, 140–141, 143–148
  backpropagation, 147
  feed-forward network, 144
  gradient descent, 147–148
  hyperparameters, 146
  learning rate, 141, 148
  loss (cost) function, 147
    categorial cross-entropy, 147
    margin (ranking) loss, 147
  multi-layer perceptron, 145
  nonlinear activation functions, 145
  parameters, 146
  perceptron, 145
  softmax function, 146
  weight matrix, 144

bag-of-words model, 44, 48, 115, 135
  continuous bag-of-words, 155, 296
black box representation, *see* distributed representation
Bloomfield, Leonard, 5–7

C-PHRASE, 163
categorical representation, 217
co-occurrence matrix, 48–50, 97
  pair-by-pattern, 106, 112
  term-document, *see* word-by-document
  word-by-document, 19, 50, 101, 113
  word-by-word, 50, 111, 123
co-occurrence weighting function, 50–59
  co-occurrence frequency, 50
  conditional probability, 52
  dynamic weighting scheme, 133
  entropy, 57–59
  log-likelihood ratios, 104
  logarithmic transformation, 52
  Pearson correlation, 100
  pointwise mutual information (PMI), 53–55
  positive local mutual information (PLMI), 55
  positive pointwise mutual information (PPMI), 54

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-107-00429-0 — Distributional Semantics
Alessandro Lenci , Magnus Sahlgren
Index
More Information

t-test, 55
tf-idf, 57–59
co-occurrences, *see* linguistic context
coercion, 293, 345
logical metonymy, 293, 345–347
probabilistic model, 345
thematic fit model, 346
selectional polysemy, 293
cognitive science, 18, 270–282,
behaviorist psychology, 14
blindness, 275
cognitive psychology, 15
concept, 18, 270–282
exemplar model, 324, 335
prototype model, 324, 341
embodied cognition, 271
embodiment, *see* grounding
estimating semantic similarity, 270
extrapolating normed data, 270
functional Magnetic Resonance Imaging (fMRI), 188, 279–281
generalized event knowledge, 339, 343, 345
grounding, 22
Chinese room argument, 262
cline of embodiment, 276
direct and indirect grounding, 274
linguistic embodiment hypothesis, 272
propagation of grounding, 276
reference problem, 268–269
symbol grounding problem, 262, 359
Symbol Interdependency Hypothesis, 273
ungrounded representation, 262
linguistic and sensory-motor experience, 18, 272–276
neurosemantic decoding, 279–282
representational pluralism, 272
semantic memory, 18, 270–282
semantic priming, 187
Syntactic Bootstrapping Hypothesis, 274
word associations, 14, 188
cognitive semantics, 22
collocation, 13, 53
compositional distributional semantics, 285–353
extrinsic evaluation, 316–317
natural language inference (NLI), 316
intrinsic evaluation, 314–317
cross-level semantic similarity, 315–317
phrase and sentence similarity, 315
compositional DSM, 294–314
Distributional Functional Model (DFM), 302–307, 348
distributional function, 302
linear regression, 305
Matrix-Vector Recursive Neural Network (MV-RNN), 307–308, 348
Practical Lexical Function (PLF), 306–307
vector composition functions, 294–300
compositionality, 285, 347–353
co-compositionality, 292
enriched composition, 293
Fregean Compositionality, *see* Principle of Compositionality
Principle of Compositionality, 286, 291, 302, 305, 308, 318, 347–353
Conceptual Semantics, 215, 290
conceptual space, 218
connectionism, *see* artificial neural network
connotative meaning, 251
context-sensitive distributional representation, 318–339
contextual DSM, 325–339
exemplar DSM, 323–325
lexical substitution, 320
vector contextualization, 320–323
word similarity in context, 319–320
contextual DSM, 96, 325–339, 351
BabyBERTa, 337
BART, 357
Bidirectional Encoder Representations from Transformers (BERT), 331–333
BERTology, 338
CLIP, 334
Contextual Vector (CoVe), 326
DALL·E, 360
DeBERTa, 334
DistilBERT, 337
ELECTRA, 334
Embeddings from Language Models (ELMo), 326–327
Generative Pre-trained Transformer (GPT), 330, 357
LaMDA, 361
mBERT, 334
Megatron-Turing NLG, 357
PaLM, 357
RoBERTa, 334
T5, 357
ViLBERT, 334

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-107-00429-0 — Distributional Semantics
Alessandro Lenci , Magnus Sahlgren
Index
More Information

XLM, 334
XLNet, 334
contextual hypothesis, *see* Distributional Hypothesis
contextual representation, 16
contextual view of meaning, 13, 15
continuous representation, 219
Convolutional Neural Network (CNN), 266, 310–311
corpus
balanced corpus, 32
corpus annotation, 34–38
chunking, 37
lemmatization, 36
part-of-speech (POS) tagging, 36
syntactic parsing, 37–38
tokenization and normalization, 34–35
corpus selection, 31–34
corpus size, 32–34
corpus vocabulary, 38
diachronic corpus, 258
frequency distribution, 27–31
parallel and comparable corpus, 244
corpus linguistics, 12–14
cross-lingual DSM, 243–251
cross-lingual contextual DSM, 334
bilingual lexicon induction, 244
cross-lingual semantic similarity, 244
joint models, 248–251
BilBOWA, 249
BiSkip, 250
cross-lingual co-occurrence matrix, 248
cross-lingual neural DSM, 249–250
cross-lingual SGNS, 249–250
multilingual Topic Model, 248
pseudo-bilingual corpus, 250
Trans-gram, 250
mapping models, 244–248
canonical correlation analysis (CCA), 247
Cross-domain Similarity Local Scaling (CSLS), 246
orthogonal transformation, 246
regression, 245–246
seed bilingual lexicon, 247
cultural stereotype, 253
gender bias, 254
curse of dimensionality, 64
hubness, 64, 246
data sparseness, 30, 64–65, 105
datasets
Almuhareb-Poesio (AP), 181
BATS, 186
Battig (Battig-2010), 181
Bicknell et al. (2010) (Bicknell), 344
BIG-bench, 362
BLESS, 186
BrainBench, 188
CogniVal, 188
CoInCo, 320
CoNLL-2003 shared task, 194
CoSimLex, 320
DiffVec, 187
DTFit, 344
ESL, 177
ESPGame, 264
ESSLLI-2008, 181
EVALution (EVAL), 186
FAST, 188
Ferretti et al. (2001) (Ferretti), 344
GLUE, 317
GOOGLE, 182
Grefenstette and Sadrzadeh (2011) (GS2011), 315
HyperLex, 187
ImageNet, 266
Kartsaklis and Sadrzadeh (2014) (KS2014), 315
McRae et al. (1998) (McRae), 344
MEN, 180
Miller and Charles (MC), 178
Mitchell and Lapata (2008) (ML2008), 315
Mitchell and Lapata (2010) (ML2010), 315
Movie Reviews (MR), 194
MSR, 182
MTurk-287 (MTurk-771), 180
MultiNLI, 317
oLMpics, 338
Padó (2007) (Padó), 344
Rare Words (RW), 178
RELPRON, 315
Rubenstein and Goodenough (RG65), 178
SAT, 182
SCWS, 319
Semantic Priming Project, 188
SemEval 2007, 320
SemEval-2 2010, task 8, 194
SemEval-2012, 187
SentEval, 317
SICK, 316

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-107-00429-0 — Distributional Semantics
Alessandro Lenci , Magnus Sahlgren
Index
More Information

SimLex-999 (SL-999), 179
SimVerb-3500 (SV-3500), 180
SNLI, 195, 317
SP-10K, 344
Stanford IMDB, 195
SuperGLUE, 317
SWAG, 317
TOEFL, 176
TR9856, 180
Turney (2012) (T2012), 315
TWSI, 320
USF, 188
USim, 319
Visa, 268
Visual Genome, 264
Wang et al. (2018) (Wang2018), 344
WiC, 320
WinoGrande, 317
WordRep, 185
WordSim-353 (WS-353), 180
decompositional (componential) semantic model, 215–216
deep learning, *see* artificial neural network
denotative meaning, 251
diachronic distributional semantics, 256–261
diachronic SGNS, 259
diachronic vector (embedding), 258
global measure of semantic change, 258
laws of semantic change, 261
local neighborhood distance, 258
offline (mapping) methods, 259–260
online methods, 259
semantic density, 260
Temporal Random Indexing (TRI), 259
vector alignment, 259–260
word sense change, 260–261
Topic Models (TMs), 260
dimensionality reduction, 65
Non-negative Matrix Factorization (NMF), 76–77
Non-negative Tensor Factorization (NTF), 342
Principle Component Analysis (PCA), 74–76
Hellinger PCA, 76
Random Projection, 128–130, 134–135
Singular Value Decomposition (SVD), 21, 67–73
inverse truncated SVD, 73, 337
truncated SVD, 69
Discourse Representation Theory, 215, 352
distant supervision, 253
distributed representation, 219–221, 349–350
Distributional Hypothesis, 3–16, 26
Strong Distributional Hypothesis, 18, 254, 271–276
Weak Distributional Hypothesis, 16, 271
distributional representation, 26, 62
context-sensitive representation, 318–339
grounded distributional representation, 261–270
distributional semantic model (DSM), 91–96
contextual DSM, 96, 325–339, 351
count models, 92
matrix models, 92, 97–125, 190
random encoding models, 94, 126–142, 191
cross-lingual DSM, 243–251
document DSM, 101, 113, 132, 192
exemplar DSM, 323–325, 333
hybrid DSM, 352
multi-prototype DSM, 225–226, 323, 336
multimodal DSM, 264–270
predict models, 96, 143–168, 191
static DSM, 96–168, 333, 338
syntax-based DSM, 103, 163, 191
window-based DSM, 98, 99, 132, 157, 191
distributional similarity, 77
distributional vector, 26, 61–67
explicit vector, 63–64
holistic vector, 185, 284, 301
implicit vector, *see* embedding
instability, 212
phrase and sentence vector, 285
prototype vector, 340
substitute vector, 324
token vector, 96, 225, 324
type vector, 63, 223, 336
word-in-context vector, 320
distributionalism, 4–10
DSM evaluation, 171–189
extrinsic evaluation, 171, 188–189, 193–195
document sentiment classification, 195
natural language inference (NLI), 195
semantic relation classification, 194
sentence sentiment classification, 194
sequence labeling (POS tagging, chunking, and NER), 194
subjectivity classification, 195
intrinsic evaluation, 171, 176–188, 193
analogy tests, 182–186

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-107-00429-0 — Distributional Semantics
Alessandro Lenci , Magnus Sahlgren
Index
More Information

categorization tests, 181–182
psycholinguistic tasks, 187–188
relation tests, 186–187
similarity and relatedness tests, 178–181
synonym tests, 176–178

efficiency, 94
ELIZA, 361
embedding
anisotropy, 336–337
pretrained embedding, 22, 154, 189, 333
sentence embedding, 308–309, 348, 351
word embedding, 21, 26, 64–67
contextual embedding, 21, 96, 226, 325–339
cross-lingual contextual embedding, 334
cross-lingual embedding, 243–251
diachronic embedding, 256–261
multimodal contextual embedding, 334
multimodal embedding, 264–270
static embedding, 96–168, 325–335
encoder-decoder model, 311–314, 327
explainability, 218
Extended Distributional Hypothesis, *see* Latent Relational Analysis

fairness, *see* word embedding debiasing
fast mapping, 267
feature, 22, 61
feature extraction, *see* dimensionality reduction
latent feature (dimension), 65–68, 113, 121
Firth, John R., 12–14
formal (model-theoretic) semantics, 23, 287–290, 352
denotation, *see* reference
extension, 234, 288
intension, 235, 288
logical form, 288, 352
model, 288
reference, 23
truth-conditions, 23, 287
Frame Semantics, 217, 290
Frege, Gottlob, 23, 286
function-argument structure, 286
function, 287, 350
function application, 287, 347
typed function application, 289
variable, 287, 350

Garvin, Paul L., 3, 19
Gated Recurrent Unit (GRU) network, 312
Generative Lexicon, 215, 293, 341, 345
Qualia Structure, 216, 345
grounded distributional representation, 261–270
denotational similarity, 264
experiential data, 263
perceptual vector, 263
Topic Models (TMs), 263

Harris, Zellig S., 4–8, 13–16, 19–21, 23
Hearst pattern, *see* lexico-syntactic pattern
hierarchical softmax, 154, 160
holistic representation, 221
hybrid DSM, 352
Structured Distributional Model (SDM), 352
hypernymy detection and prediction, 232–239
directional similarity measure, 233–235
APinc, 234
CL, 234
invCL, 234
WeedsPrec, 234
Distributional Inclusion Hypothesis (DIH), 233
graded entailment, 232
hypernymy-specific embedding, 237–239
HyperScore, 239
LEAR, 237
pattern-based models, 232–233
semantic generality, 235–236
SLQS, 235
supervised models, 236–237
lexical memorization, 237
vector combination, 236

incrementality, 94, 134, 365
information retrieval, 57, 173
Latent Semantic Indexing (LSI), *see* Latent Semantic Analysis (LSA)
Vector Space Model, 19–20, 48, 92, 97, 100, 113
interpretability, 218, 276, 337

Jakobson, Roman, 8–9
Johnson-Lindenstrauss lemma, 126–128

Kanerva, Pentti, 130

landmark method, 321

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-107-00429-0 — Distributional Semantics
Alessandro Lenci , Magnus Sahlgren
Index
More Information

latent semantic space, *see* latent feature (dimension)
lexical ambiguity, 221
  homonymy, 221
  polysemy, 221, 293, 335
lexical entailment, *see* hypernymy
lexical semantic competence, 214
  inferential competence, 261
  referential competence, 261
lexico-syntactic pattern, 105, 232
linear algebra, 20
  cosine, 82
  covariance matrix, 74
  eigenvalue, 74
  eigenvector, 74
  linear transformation (map), 74
    orthogonal transformation, 246
  matrix, 49
    factorization, 67
    orthogonal, 246
    product, 68
    transpose, 68
  orthogonal decomposition, 83
  Procrustes problem, *see* orthogonal transformation
  tensor, 108, 303
    tensor dimensionality reduction, 110
    matricization, 110–111
  vector, 49
    addition (sum), 294
    centroid, 224
    circular convolution, 136, 299
    circular correlation, 137
    difference, 123, 183–184
    dot (inner) product, 68
    elementwise multiplication, 297
    length (Euclidean norm), 61
    normalization, 81
    parallelogram law, 183
    projection, 83
    tensor product, 298
    vector-matrix multiplication, 74, 145
  vector space, 61
linguistic context, 5, 38–48, 91–92
  collocate, 13, 39
    syntactic collocate, 44–46
    window-based collocate, 40–44
  construction, 46–47
  context lexeme, 39
  context selection, 59–60
  context window, 40–43
  document, 47–48
linguistic distribution, 5–8, 38
Long Short-Term Memory network (LSTM), 311–312
  biLSTM, 326

meaning postulate, 216
metaphor, 292
metonymy, 292
Miller, George A., 15–16
Montague Semantics, 288–289, 302–304
  rule-by-rule interpretation, 289
  semantic type, 288
  typed function application, 289
Montague, Richard, 23, 288
multi-prototype DSM, 225–226, 323, 336
  multi-prototype similarity measure, 226
multimodal DSM, 264–270
  multimodal contextual DSM, 334
  cross-modal mapping, 268
  grounded visual attribute, 268
  image dispersion, 276
  image vector, 266
    Bag-of-Visual-Words (BoVW), 266
    Convolutional Neural Network (CNN), 266, 310–311
  Multimodal Skip-Gram (MMSG), 267
  multimodal vector (embedding), 265
  visual semantic generality, 236
multimodal representation, 220
multiword expressions, 35, 300–302
  compositionality prediction, 300–302

nearest neighbor, 78
  paradigmatic neighbor, 229
  syntagmatic neighbor, 229
negation, 243
negative sampling, 160–161
neural language model, 148–156, 326–333
  autoregressive (causal) model, 330
  bidirectional model, 331
  denoising autoencoding, 268, 331
  discriminative model, 154
  masked language modeling, 331
  masked multimodal learning, 334
  multilingual masked language modeling, 334
  multimodal alignment prediction, 334
neurosemantic decoding, 279–282
  neural decoding mapping (decoder), 280
  voxel, 280

one-hot vector, 130, 147, 149

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-107-00429-0 — Distributional Semantics
Alessandro Lenci , Magnus Sahlgren
Index
More Information

paradox of simultaneous difference and similarity, *see* antonymy
Parallel Distributed Processing, *see* artificial neural network
Plato's problem, 362
probing task, 277, 317
  edge probing, 318
  SemRole, 318
productivity, 284, 347–353
  partial productivity, 351
prompting, 338, 358

relational semantic model, 216–217
reporting bias, 273, 364
representation degeneration problem, *see* anisotropy
representation learning, 148, 358
Representational Similarity Analysis (RSA), 205–206, 281
  second-order isomorphism, 206
Rescorla-Wagner model, 167

Sapir, Edward, 5
Saussure, Ferdinand de, 8–10
scalability, 94
selectional preferences, 289, 293, 321, 339–345
  clustering, 340
  exemplar model (EPP), 341
  inverse preferences, 321, 341
  multi-way selectional preferences, 342
  neural models, 343
  prototype model, 342
  pseudo-disambiguation task, 343
  selectional polysemy, 293
  thematic fit, 339, 344
  Topic Models (TMs), 342
selectional restrictions, *see* selectional preferences
self-similarity measure, 335
self-supervised learning, 149, 326
semantic change, *see* diachronic distributional semantics
semantic differential, 15
semantic feature, 10, 215, 276–278
  feature (property) norm, 219, 263
  feature mapping, 277
  semantic feature vector, 218
semantic field, 10
semantic network, 216
semantic relation, 174
  detection (classification), 186, 230
  paradigmatic, 10, 19, 174, 229–243
    antonymy, 10, 174, 239–240
    co-hyponymy, 10, 174, 231
    hypernymy, 10, 174, 231–239
    synonymy, 10, 174
semantic role, 174, 216
semantic similarity, 15, 172–176
  attributional (taxonomic) similarity, 172–173
  cross-lingual similarity, 244
  relational (analogical) similarity, 105, 172–173
  semantic relatedness, 173
  thematic relatedness, 173
  topical relatedness, 173
semantic space, 15, 62
semantic specialization, 237–239, 242
semantic type, 288, 290, 339–340
sentence embedding, 308–314, 348, 351
  Convolutional Neural Network (CNN), 310–311
  encoder-decoder model, 311–314
    autoencoder, 311
    DisSent, 311
    FastSent, 311
    InferSent, 311
    Paragram, 311
    Sentence-BERT, 335
    SkipThought, 313
    Universal Sentence Encoder, 313
  Paragraph Vector (doc2vec), 309
  Sent2Vec, 310
seq2seq model, *see* encoder-decoder model
similarity, 91, 172, 350
Simple Recurrent Network (SRN), 149–151, 311
  vanishing gradient, 311
Sinclair, John M., 13
singleton fallacy, 360
sparse distributed memory, *see* Random Indexing (RI)
static DSM, 96–168, 333, 338
  Bound Encoding of the Aggregate Language Environment (BEAGLE), 135–138
  context2vec, 164
  Correlated Occurrence Analogue to Lexical Semantics (COALS), 100
  Dependency Vectors (DV), 103–104
  Distributional Memory (DM), 107–113
    weighted tuple, 108

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-107-00429-0 — Distributional Semantics
Alessandro Lenci , Magnus Sahlgren
Index
More Information

Explicit Semantic Analysis (ESA), 102
FastText, 164–165
GBOW, 163
Global Vectors (GloVe), 122–124
probability ratio, 123
weighted regression, 124
Hierarchical Log-BiLinear model (HLBL), 154
hierarchical softmax, 154
High Dimensional Explorer (HiDEx), 100
Hyperspace Analogue to Language (HAL), 99–100
Latent Relational Analysis (LRA), 104–107
Latent Semantic Analysis (LSA), 21, 100–103
Log-BiLinear model (LBL), 153
Neural Network Language Model (NNLM), 152
embedding layer, 152
Random Indexing (RI), 130–135
RI with random permutations, 138–140
Self-Organizing Maps (SOM), 140–142
Topic Models (TMs), 113–122
Bayesian inference, 113
Latent Dirichlet Allocation (LDA), 115
latent topic, 113
probabilistic Latent Semantic Analysis (pLSA), 115
topic vector, 120–122
vLBL and ivLBL, 155
noise-contrastive estimation, 155
word2vec, 156–163
Continuous Bag-Of-Words (CBOW), 158–160
dynamic window, 162
hierarchical softmax, 160
Skip-Gram (SG), 157–158
Skip-Gram with Negative Sampling (SGNS), 160–161
subsampling, 163
word2vecf, 163
stop words, 40
structural linguistics, 5–10
symbolic models, 215–218, 286–291, 349–350
syntactic dependency, 37, 44, 103
syntagmatic and paradigmatic relations, 9–10, 13–15
systematicity, 349
target lexemes, 38
thesaurus, 19–20
training corpus, *see* corpus
transfer learning, 189, 333–334
cross-lingual transfer learning, 244
fine-tuning, 22, 333, 358
pretrained embedding, 22, 154, 189, 333
pretrained feature, *see* pretrained embedding
Transformer, 327–330
attention, 327–328
multi-head self-attention, 328–330
positional embedding, 328
Turing test, 361

usage-based theory, 11, 22, 340

vagueness, 221, 292
vector (dis)similarity measure, 77–87
cosine similarity, 81
Dice similarity, 85
Euclidean distance, 80
Jaccard measure, 85
Jensen-Shannon (JS) divergence, 87
Kullback-Leibler (KL) divergence, 86
Lin measure, 85
Manhattan distance, 80
vector composition functions, 294–300, 348
additive model, 294–297
circular convolution, 298
multiplicative model, 297–298
tensor product, 298–300
vector contextualization
combination with selectional preferences, 321–322
multi-prototype DSM, 323
Topic Models (TMs), 322
vector component re-weighting, 322
word-in-context vector, 320

weight-tying trick, 332, 337
Wittgenstein, Ludwig, 11–12
language game, 12
word embedding bias, 254–255
Word-Embedding Association Test (WEAT), 254
word embedding debiasing, 255–256
learning debiased embeddings, 256
post-processing approach, 255
word sense, 221–229, 335
context clustering, 224–227
context vector, 224

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-107-00429-0 — Distributional Semantics
Alessandro Lenci , Magnus Sahlgren
Index
More Information

multi-prototype DSM, 225–226, 323, 336
sense vector (embedding), 224
neighbor clustering, 227–229
Clustering by Committee (CBC), 227
graph-based clustering, 228
syntagmatically labeled partitioning, 228
sense cluster, 224
Topic Models (TMs), 226
word sense disambiguation (WSD), 222
word sense induction (WSI), 223
WordNet, 216

zero-shot learning, 245, 269, 278, 280, 338, 358
Zipf's law, 28–31

© in this web service Cambridge University Press & Assessment www.cambridge.org