MUME 2016 - The Fourth International Workshop on Musical Metacreation, ISBN #978-0-86491-397-5

Learning Musical Creativity via Stochastic Transduction Grammars: Combination, Exploration and Transformation

Dekai Wu
Human Language Technology Center
Department of Computer Science and Engineering
Hong Kong University of Science and Technology
dekai@cs.ust.hk

Abstract

We discuss how Boden's creative processes of combination, exploration, and transformation naturally emerge in models that learn musical improvisation via stochastic transduction grammar induction. Unlike a conventional monolingual grammar, a transduction grammar represents complex transformative relationships between one representation language and another. For musical improvisation, a transduction grammar both provides a large (typically infinite) space of possible hierarchical combinations, and defines a combinatorial space to explore. A stochastic transduction grammar (STG) allows controlled randomness in the combination and exploration. We have been developing STG based models in recent work on learning musical improvisation for hip hop, flamenco, and blues. Inducing an STG simultaneously (a) identifies chunks that will become candidates for recombination as well as patterns of combination, (b) constructs a new spaces for exploration in improvisation and composition, and (c) learns transformations from one representation to another.

Introduction

We are developing a general computational framework for learning musical improvisation, that (a) is capable of representing a realistically broad range of the many different complex interactions among factors that should influence the improvisation, and yet (b) can still support efficient polynomial-time training and improvisation algorithms. While a full solution to the representation, learning, and improvisation problems is clearly a long-term research program, we have already begun to show how various aspects of these tasks can be accomplished, via bilingual **stochastic transduction grammar** or **STG** models that can simultaneously capture contextual preferences across a wide variety of dimensions.

In this paper, we discuss why we believe that our STG modeling framework tackles the classic modes of creativity advanced by Boden (1994, 2004, 2009). While numerous elaborations, modifications, and criticisms of Boden's analysis have since been proposed, her breakdown of creative processes into unfamiliar combinations of familiar concepts, exploration of conceptual spaces, and transformation of conceptual spaces still does identify some of the major essential cornerstones of creativity. Boden's first type of creativity, through novel combinations of familiar concepts, leads to the kind of surprise that comes from unfamiliarity or unlikeliness. Boden's second type of creativity, through exploration of an extremely large or infinite conceptual space, leads to the kind of surprise that comes from seeing that an unexpected idea actually fits within your existing conceptual framework, but you hadn't thought of it. Boden's third type of creativity, through transformation of ideas represented in one conceptual space to another, leads to the kind of surprise that comes from seeing something that could never have appeared to be possible because of the limitations of the original conceptual space.

Musical improvisation is a particularly interesting domain to study creativity, being a uniquely human province of creative expression that has not been found even in other "singing" species. Musical improvisation can be seen as the creative activity of spontaneous, on-the-fly musical composition without prior planning, in response to a novel context, in contextually relevant ways that adhere to stylistic conventions, yet are not constrained by *a priori* written scores. The novel context can be (a) provided by other musicians, who are often also improvising, or (b) provided simply by the environment, if a musician is producing a solo improvisation, and in the extreme case could even be an "empty" null environment. Whereas Western music in recent centuries has placed a premium on musicians executing pre-written scores, historically speaking this is a recent anomaly—throughout most of human history, instead it is creative improvisation that has been the norm in many, if not most, traditional and folk forms of music.

We identify five aspects of the STG modeling framework that bear upon Boden's analysis:

1. Our shift to *stochastic* rather than purely symbolic grammars in the late 1980s and early 1990s (Wu, 1989, 1991, 1993, 1995), which have since become dominant in most relevant research strands, emphasize probabilistic judgments of the degree of goodness for various hypotheses about possible combinations, as well as contextually informed search biases in exploration.

The Fourth International Workshop on Musical Metacreation, MUME 2016.
www.musicalmetacreation.org.

2. Our emphasis on hierarchically *compositional* models rather than flat models (like Markov models) provides a sufficiently rich space of combination.
3. The shift of the primary emphasis to bilingual modeling of structural *transduction* shifts the brunt of the modeling power onto learning biases in the *transformations* between alternative conceptual representations, rather than traditionally limited monolingual modeling of a single representation language.
4. Our emphasis on *induction* methods for STGs directly attacks the problem of generating new conceptual spaces to explore, since exposure to different environments automatically results in inducing different spaces of possible transformations.
5. Our *transduction neural network* implementations of STGs further increase computational efficiency while providing a parallel attack on the combinatorial explosion of category induction when generating new conceptual spaces.

We discuss these relationships to Boden's characterization of creativity in the context of our recent work deploying stochastic transduction grammars in hip hop, flamenco, and blues learning models. In the hip hop improvisation models of Wu et al. (2013), we showed how STGs can be used to learn how to improvise responses in freestyle rap battling when confronted with arbitrary challenge raps, by learning complex relationships between challenges and responses. In the flamenco learning models of Wu (2013), we showed how STGs can be used to learn how to improvise complementary lines in, for example, *palmas* percussion in the context of perceiving *cajn* percussion, by learning complex hypermeter and rhythm biases in the relationship between the languages of different percussion instruments. In the blues learning models of Wu (2016), we showed how STGs can be use to learn the degree to which microtonal "blues notes" tend to be bent depending on the context of improvisation.

Formal preliminaries: Transduction grammars

Musical improvisation modeling approaches based on STGs benefit from leveraging several decades of advances in the field of statistical machine translation, which exhibits very analogous challenges. Transduction grammars represent relationships between alternate representations of a complex idea, via a *bilingual* parse tree. The syntax directed transduction grammars (SDTGs) of classic formal language theory (Lewis and Stearns, 1968) are powerful, but exponentially too complex for recognition, parameter estimation, and induction. Our work has pioneered the learning and use of a strongly restricted subclass of SDTGs known as inversion transduction grammars or ITGs (Wu, 1995, 1997). Unlike SDTGs, polynomial time training and improvisation algorithms exist for ITGs, along with their family of restricted variants, including linear transduction grammars (LTGs) and monotonic transduction grammars (MTGs). What makes ITGs of even more special interest is that they have been empirically demonstrated over several decades of cross-lingual and machine translation research to possess nearly universal coverage of transformations between any pair of natural language representations (Zens and Ney, 2003; Saers, Nivre, and Wu, 2009; Addanki et al., 2012). The combinatorial properties of ITGs explain a longstanding mystery in linguistics of the "magic number 4" language universal of semantic frame structure, that suggests a deep-rooted evolutionary reason why human cognition has evolved to only be capable of efficiently transforming certain classes of representation languages (Wu, 2014).

Transduction rules (and instances of rules) represent structured correlations between an input representation language and an output representation language. Formally, an ITG is a tuple $\langle N, \Sigma, \Delta, R, S \rangle$, where N is a finite nonempty set of nonterminal symbols, Σ is a finite set of terminal symbols in L_0 (output language), Δ is a finite set of terminal symbols in L_1 (input language), R is a finite nonempty set of inversion transduction rules and $S \in N$ is a designated start symbol. A normal-form ITG consists of rules in one of the following four forms:

$$S \to A, A \to [BC], A \to \langle BC \rangle, A \to \mathbf{e}/\mathbf{f}$$

where $S \in N$ is the start symbol, $A, B, C \in N$ are *nonterminal* symbols and $\mathbf{e}/\mathbf{f}$ is a *biterminal*. A biterminal is a pair of symbol sequences: $\Sigma^* \times \Delta^*$, where at least one of the sequences have to be nonempty. The square and angled brackets signal straight and inverted order respectively. With straight order, both the L_0 and the L_1 productions are generated left-to-right, but with inverted order, the L_1 production is generated right-to-left.

Given a pair of input and output sentences $e_1, \ldots, e_T$ and $f_1, \ldots, f_V$ respectively, an ITG generates a biparse tree by recursively combining smaller bispans (chunks of aligned input and output segments) into larger bispans using the syntactic rules in straight or inverted order. Each bispan corresponds to at least one nonterminal and is represented using a 4-tuple s, t, u, v which corresponds to the input segment with tokens $e_s, e_{s+1}, \ldots, e_t$ and the output segment with tokens $f_u, f_{u+1}, \ldots, f_v$.

Sets of biparse trees are represented explicitly as well, but for efficiency, tabular and hypergraph data structures are used wherever possible to compress the storage of biparse trees that share subtrees (these data structures are commonly referred to as charts or packed forests). This is responsible for the polynomial-time dynamic programming algorithms for recognition and training for ITGs.

Stochastic modeling

One of the key innovations of Wu (1997) was to introduce *stochastic* versions of transduction grammars. Boden's analysis of creativity, developed largely before the wholesale paradigm shift of AI to statistical learning approaches, did not benefit from the additional mileage afforded by probabilistic approaches.

In STGs, such as stochastic ITGs, a probability is associated with each transduction rule. Typically, this is the conditional probability of the right-hand-side, given the left-hand-side terminal. This corresponds to a recursive stochastic generative process, in which the left-hand-side nonterminal probabilistically generates a combination of elements (either nonterminals or biterminals) described by the right-hand-side. Stochastic generative models facilitate controlled randomness in all of Boden's types of creativity.

For combination and transformation processes, stochastic transduction rules dictate not only what types of ideas may be combined in improvisation, but also how good such a combination is deemed to be. Of course, the probability of the elements being combined also contributes to the overall degree of goodness, given the recursive tree-structured compositionality. The rap battle learning approach introduced by Wu, Addanki, and Saers (2013) improvises freestyle rap responses by considering alternative ways to combine ideas that are somehow relevant to the challenge rap. Combination and transformation processes are integrated in the STG model—the combination hypotheses arise in the course of transforming the challenge rap into a response rap. The choice of the complete response follows the distribution over the space of transductions implied by the transduction rule probabilities.

For exploration processes, the probabilities of hypotheses (kept in the aforementioned hypergraphs) facilitate biasing the directions of exploration toward more likely directions. Because of the large number of choices at each level of granularity, it is typical to use various pruning heuristics to limit the number of improvisation hypotheses at each level. This can also be thought of as using the environmentally trained rule probabilities to dynamically reshape the conceptual space. Exploration decisions can be made either by simply retaining the *n* best hypotheses at each level, or if it is desired to increase the "surprise factor", by selecting a subset of the hypotheses randomly according to the distribution.

Compositional modeling

In previous work on stochastic grammatical models for music, it is common to find flat Markov models and/or hidden Markov models (HMMs). For example, both the Continuator model of Pachet (2003) and the Factor Oracle models of Assayag et al. (2006) and Assayag and Dubnov (2004) use Markov models to learn music improvisation conventions—an approach further explored by Franois, Chew, and Thurmond (2007) and Franois, Schankler, and Chew (2010). A grammar induction approach for learning jazz grammars under Markovian assumptions is proposed by Gillick, Tang, and Keller (2010). Relatively little has been done on musical structure modeling using stochastic context-free grammars (Lari and Young, 1990). Related work on unsupervised learning of CCMs (a variant of SCFGs) for musical grammars includes that of Swanson, Chew, and Gordon (2007), or the Data Oriented Parsing approach of Bod (2001).

From the standpoint of Boden's combination processes, a general framework for creative improvisation needs a less restricted mechanism for creating unfamiliar combinations by combining familiar ideas. Such combination cannot be limited to flat models. As Boden (1994) writes:

> For writing [chord] sequences, unless they are kept boringly simple, typically requires a great deal of time and effort. They are complex hierarchical structures, with subsections "nested" at several different levels, and with complex harmonic constraints linking sometimes far-separated chords. (p. 91-2)

Our STG framework moves strongly toward hierarchically *compositional* modeling. This allows such contexts for improvisation to be represented adequately—even though Boden suggests that such chord progressions could not be improvised "on the fIy", skilled jazz musicians can in fact easily improvise chord progressions of the same complexity as most jazz standards, so we do not want a representation that precludes describing such conceptual combinations.

Transduction oriented modeling

The shift of emphasis to *bilingual* transduction grammars can be seen as a direct formalization of Boden's characterization of the role of *transformation* in creativity. It recognizes that realistically human musical improvisation must require not only complex combinations of structures and patterns, but also finely tuned predictions about how to interpret an environmental context represented in an input conceptual space, and transform it into an improvision in an output conceptual space. Improvisational and accompaniment decisions in one part can be influenced strongly, or subtly, by decisions made in other parts; the interaction occurs hierarchically or *compositionally* at many overlapping levels of granularity. Improvisation and accompaniment decisions are not merely random; skilled participants understand how to communicate with each other within accepted conventions and frameworks—as, for example, in flamenco *palos*, Indian *ragas*, jazz and blues. Widely used conventions include tonal systems, metrical constraints, chord progressions, verse structures, rhythmic patterns, and melodic phrases that are re-used or swapped into different positions within the structures.

A good model for musical improvisation should allow making decisions that integrate interacting contextual factors over many levels of granularity, that is capable of encoding such sophisticated phenomena. However, the longstanding risk in using expressive representations for transduction is the typically exponential blow up in the complexity of machine learning.

Stochastic inversion transduction grammars (a) have sufficient expressiveness to represent compositionally interacting factors between two different parts or instruments at many overlapping levels of granularity, (b) can be efficiently induced via the polynomial-time learning algorithms that exploit the combinatorial structure of SITGs, and (c) can then use the learned knowledge representation to creatively perform real-time improvisational expression. For capturing the complexity of hierarchical structural relationships between different musical languages, the bilingual approaches of STGs have many appealing properties. As shown in the machine translation work discussed earlier, they allow idiomatic constructs of significant complexity to be encoded. They allow biasing of probabilities from many different contextual features. They allow idiomatic constructs to be combined in creative new ways inspired by the unplanned contextual factors. They accommodate correlations that are not necessarily aligned in time, which make them significantly more expressive than context-free grammars (CFGs)—a flexibility that is exploited in both the hip hop improvisation models of Wu et al. (2013) and the blues note learning models of of Wu (2016). The basic time complexities for SITG recognition and training are $O\left(n^6\right)$, in contrast to $O\left(n^3\right)$ for stochastic CFGs, but this has still proven quite feasible especially with standard heuristic beam pruning methods.

Transduction grammar induction modeling

It can be relatively easy to construct automatic music generation algorithms that can be parametrized by various conditions and constraints. Early approaches to musical improvisation modeling often relied on manually constructed rules; these approaches can represent fairly complex kinds of structures and patterns, but the improvisation is limited to the rules that have been imagined by experts and hand coded in advance, which can only crudely be matched to true human improvisation. STGs can also be manually written if desired, to build such systems. It has been frequently argued, however, that such systems are not truly creative.

On the other hand, we believe truly creative improvisation requires learning—human improvisation arises from adapting and synthesizing experience gained in many previously unrelated scenarios. However, previous machine learning approaches to musical improvisation attempt to match their performance more finely to human improvisation by training contextual predictors on actual music data, but improvisation tends to be restricted to what can be modeled via fairly simple representations such as HMMs to limit the complexity of the learning.

Truly modeling the complexity of human musical improvisation requires machine learning approaches that not only estimate the probabilities in some existing structural model (say, an STG), but can in fact learn the *structure* of a new STG by inducing new types of combination and transformation patterns (transduction rules) and categories (nonterminals).

We have developed numerous transduction grammar induction methods over the past decades in machine translation research. Some employ bottom-up incremental rule chunking; others employ top-down incremental rule segmentation. Some induction methods are driven by a maximum likelihood objective; others are driven by a minimum description length (MDL) or maximum *a posteriori* (MAP) objective. The flamenco learning model of Wu (2013), for example, employs a MDL driven top-down incremental rule segmentation strategy for learning transduction rules and inducing categories.

Inducing an STG inherently defines a new conceptual space for Boden's exploration processes. Newly induced categories and rules dictate a space of possible combinations to explore. Transduction grammar induction changes the shape of the exploration space. It identifies chunks that will become candidates for recombination as well as patterns of combination, (b) constructs a new spaces for exploration

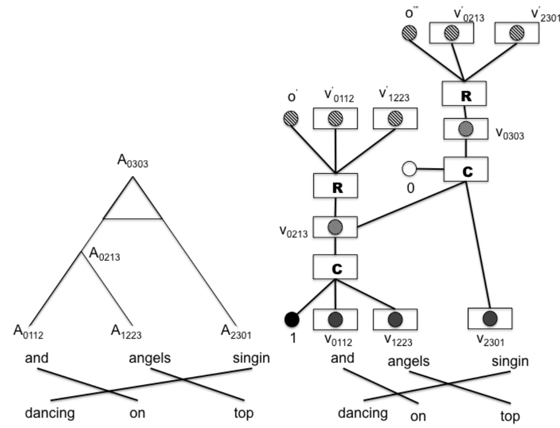

Figure 1: A symbolic biparse tree (left) and its implementation in a TRAAM neural network (right). The horizontal bar under the root node of the symbolic biparse tree indicates an inverted transduction rule at that level.

in improvisation and composition, and (c) learns transformations from one representation to another.

Transduction neural network modeling

Stochastic transduction grammars can be implemented neurally instead of symbolically, yielding tradeoffs that include a number of advantages. The TRAAM (transduction RAAM) model proposed by Addanki and Wu (2014) reduces the need to explicitly represent enormous numbers of similar competing hypotheses, by instead representing improvisation hypotheses using fixed-dimensionality continuous vectors. The distributed vector representations in TRAAM aim to parallel the structural composition of a syntax directed transduction grammar. However, unlike symbolic transduction grammar based representations, the continuous vector representations in effect represent soft neighborhoods of cross-lingual transformational associations. TRAAM implicitly learns context-sensitive generalizations over the structural relationships, between the corresponding parts of the input and output representations across all levels of granularity, while avoiding incurring the symbolic models' exponential cost of modeling context sensitivity.

More formally, TRAAM is a bilingual generalization of the way that the RAAM (recursive autoassociative memory model) of Pollack (1990) monolingually approximates context-free grammars. In TRAAM's distributed representation of an ITG, each bispan s, t, u, v is represented using a feature vector $\mathbf{v}_{stuv}$ of dimension $d$which represents a fuzzy encoding of all the nonterminals that could generate the bispan. This stands in contrast to the symbolic ITG where each nonterminal that generates the bispan must be enumerated separately. As with symbolic ITGs, vectors corresponding to larger bispans are recursively generated from the vectors representing smaller bispans, but in TRAAM this is done using a *compressor* network. The compressor network takes two vectors of dimension d, along with a single bit corresponding to straight or inverted order, and outputs a vector of dimension d—essentially compressing an input of $2\mathrm{d}+1$ dimensions to a vector of dimension d.

The role of the compressor network is analogous to the transduction rules in the ITG model, but with the important distinction of (1) keeping the encoding fuzzy, and (2) forcing generalization over similar vectors in the Euclidean space neighborhood. Using an example from the neural hip hop improvisation work of Wu and Addanki (2015), figure 1 visualizes how transduction rule instances (both straight and inverted) correspond to inputs to the compressor network. Each nonterminal in an ITG can be encoded as a bit vector, identical to the vector of the bispan in our model. Using the universal approximation theorem of neural networks (Hornik, Stinchcombe, and White, 1989), an encoder with a single hidden layer can represent any set of transduction rules. Conversely, any variant of our model can be represented as an ITG by assuming a unique nonterminal label for the feature vector corresponding to each bispan.

Hence, Boden's transformations between conceptual spaces can be represented using neural TRAAMs as an alternative to symbolic STGs, providing two ways of encoding compositional bilingual relations that can model the novel combination of familiar concepts. TRAAM's neural encoding of nonterminals is better suited for modeling generalizations over bilingual relations without exploding the space for Boden's exploration processes, while symbolic ITG representations avoid potential confusions between concepts due to accidental similarities between vectors.

Conclusion

Musical improvisation in our stochastic transduction grammar based framework is modeled as a quasi-translation task in which any environmental context encoded in an input representation language is probabilistically transformed into a potentially novel, unfamiliar combination of familiar ideas in an output representation language. This is translation, or *transduction*, in the mathematical sense, as in formal language theory; it is of course not translation in the linguistic sense. We have suggested how our STG approach addresses each of Boden's major types of creative processes: combination, exploration, and transformation. Our shift to *stochastic* rather than purely symbolic grammars emphasizes probabilistic judgments of the degree of goodness for (a) alternative transformation hypotheses that are created from novel combinations of familiar ideas, as well as (b) contextually informed search biases in exploration. Our emphasis on hierarchically *compositional* models provides a richer, more realistic space of combination than flat Markov models. Our shift to emphasize bilingual modeling of structural *transduction* shifts the modeling focus onto learning biases in the *transformations* between alternative conceptual representations, instead of limiting ourselves to traditional monolingual modeling of a single representation language. Our emphasis on *induction* methods for STGs directly attacks the problem of generating new conceptual spaces to explore, since exposure to different environments automatically results in inducing different spaces of possible transformations. Finally, our *transduction neural network* implementation of STGs reduces the combinatorial explosion problems of category induction when generating new conceptual spaces.

Acknowledgements

This material based upon work supported in part by Hong Kong Research Grants Council (RGC) research grants GRF16210714, GRF16214315, GRF620811 and GRF621008; the Defense Advanced Research Projects Agency (DARPA) under LORELEI contract HR0011-15-C-0114, BOLT contracts HR0011-12-C-0014 and HR0011-12-C-0016, and GALE contracts HR0011-06-C-0022 and HR0011-06-C-0023; and by the European Union under the Horizon 2020 grant agreement 645452 (QT21) and FP7 grant agreement 287658. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of RGC, DARPA, or the EU.

References

Addanki, K., and Wu, D. 2014. Transduction Recursive Auto-Associative Memory: Learning bilingual compositional distributed vector representations of Inversion Transduction Grammars. In *Proceedings of SSST-8, Eigth Workshop on Syntax, Semantics and Structure in Statistical Translation (at EMNLP 2014)*, 112–121.

Addanki, K.; Lo, C.-K.; Saers, M.; and Wu, D. 2012. LTG vs. ITG coverage of cross lingual verb frame alternations. In *16th Annual Conference of the European Association for Machine Translation (EAMT-2012)*, 295–302.

Assayag, G., and Dubnov, S. 2004. Using factor oracles for machine improvisation. *Soft Computing* 8:1432–7643.

Assayag, G.; Bloch, G.; Chemillier, M.; Cont, A.; and Dubnov, S. 2006. OMax Brothers: A dynamic topology of agents for improvization learning. In *First ACM Workshop on Audio and Music Computing Multimedia*, 125–132.

Bod, R. 2001. Stochastic models of melodic analysis: Challenging the gestalt principles. *Journal of New Music Research* 30(3).

Boden, M. A. 1994. What is creativity? In Boden, M. A., ed., *Dimensions of Creativity*. Cambridge, MA: MIT Press. 75–117.

Boden, M. A. 2004. Creativity in a nutshell. In Boden, M. A., ed., *The Creative Mind: Myths and Mechanisms*. London: Routledge. 1–10.

Boden, M. A. 2009. Computer models of creativity. *AI Magazine* 23–34.

Franois, A. R.; Chew, E.; and Thurmond, D. 2007. Mimi - a musical improvisation system that provides visual feedback to the performer. Technical Report 07-889, USC Computer Science Department.

Franois, A. R.; Schankler, I.; and Chew, E. 2010. Mimi4x: An interactive audio-visual installation for high-level structural improvisation. In *IEEE International Conference on Multimedia and Expo (ICME 2010)*, 1618–1623.

Gillick, J.; Tang, K.; and Keller, R. M. 2010. Machine learning of jazz grammars. *Computer Music Journal* 34(3):56–66.

Hornik, K.; Stinchcombe, M.; and White, H. 1989. Multilayer feedforward networks are universal approximators. *Neural networks* 2(5):359–366.

Lari, K., and Young, S. J. 1990. The estimation of stochastic context-free grammars using the inside-outside algorithm. *Computer Speech and Language* 4:35–56.

Lewis, P. M., and Stearns, R. E. 1968. Syntax-directed transduction. *Journal of the Association for Computing Machinery* 15(3):465–488.

Pachet, F. 2003. The Continuator: Musical interaction with style. *Journal of New Music Research* 32(3):33–341.

Pollack, J. B. 1990. Recursive distributed representations. *Artificial Intelligence* 46(1):77–105.

Saers, M.; Nivre, J.; and Wu, D. 2009. Learning stochastic bracketing inversion transduction grammars with a cubic time biparsing algorithms. In *11th International Conference on Parsing Technologies (IWPT'09)*, 29–32.

Swanson, R.; Chew, E.; and Gordon, A. S. 2007. Supporting musical creativity with unsupervised syntactic parsing. In *AAAI Spring Symposium on Creative Intelligent Systems*.

Wu, D.; Addanki, K.; and Saers, M. 2013. Modeling hip hop challenge-response lyrics as machine translation. In *Machine Translation Summit XIV (MT Summit 2013)*.

Wu, D., and Addanki, K. 2015. Learning to rap battle with bilingual recursive neural networks. In *24th International Joint Conference on Artificial Intelligence (IJCAI 2015)*, 2524–2530.

Wu, D.; Addanki, K.; Saers, M.; and Beloucif, M. 2013. Learning to freestyle: Hip hop challenge-response induction via transduction rule segmentation. In *2013 Conference on Empirical Methods in Natural Language Processing (EMNLP 2013)*, 102–112.

Wu, D. 1989. A probabilistic approach to marker propagation. In *11th International Joint Conference on Artificial Intelligence (IJCAI-89)*, 574–580.

Wu, D. 1991. A continuum of induction methods for learning probability distributions with generalization. In *Thirteenth Annual Conference of the Cognitive Science Society (CogSci 1991)*, 949–953.

Wu, D. 1993. Approximating maximum-entropy ratings for evidential parsing and semantic interpretation. In *13th International Joint Conference on Artificial Intelligence (IJCAI-93)*, 1290–1296.

Wu, D. 1995. Stochastic Inversion Transduction Grammars, with application to segmentation, bracketing, and alignment of parallel corpora. In *14th International Joint Conference on Artificial Intelligence (IJCAI-95)*, 1328–1334.

Wu, D. 1997. Stochastic Inversion Transduction Grammars and bilingual parsing of parallel corpora. *Computational Linguistics* 23(3):377–404.

Wu, D. 2013. Simultaneous unsupervised learning of flamenco metrical structure, hypermetrical structure, and multipart structural relations. In *14th International Society for Music Information Retrieval Conference (ISMIR 2013)*.

Wu, D. 2014. The magic number 4: Evolutionary pressures on semantic frame structure. In *10th International Conference on the Evolution of Language (Evolang X)*.

Wu, D. 2016. How blue can you get? learning structural relationships for microtones via continuous stochastic transduction grammars. In *Seventh International Conference on Computational Creativity (ICCC 2016)*.

Zens, R., and Ney, H. 2003. A comparative study on reordering constraints in statistical machine translation. In *41st Annual Meeting of the Association for Computational Linguistics (ACL-2003)*, 192–202.