
Formal Grammars for Computational Musical Analysis

The Blues and the Abstract Truth

Mark Steedman, Informatics, Edinburgh

Music Informatics and Cognition Workshop,

Sonic Arts Research Centre, Queen's University Belfast, April 2004

[When I	recall the love I	found on]	$(S/S)/NP$
[Dm7	Bm7 \flat 5 E7 \flat 9	Am7]	$(Dm7/F\sharp m7)/Dm$
[I could kiss the	ground on]		S/NP
[F \sharp m7 B7	Em7 A7]		$F\sharp m7/Dm7$
[Green Dolphin	Street]		$S\backslash(S/NP)$
[Dm7 G7]	Dm7/C7	[C]	$(C\backslash C)\backslash(Y7/C7)$

N. Washington/B. Kaper (1947) *On Green Dolphin Street*

The Jazz Blues as a Language

- There are in principal infinitely many variations on the primordial Twelve-bar Blues chord sequence that jazz musicians recognize as such, just as we all understand the infinite variety of English sentences.
- They have been explored by the likes of Louis Armstrong, Charlie Parker, and our contemporaries.

a)	I(M7)	IV(7)	I(M7)	I(7)	IV(7)	IV(7)	I(M7)	I(M7)	V(7)	V(7)	I(M7)	I(M7)
b)	I(M7)	IV(7)	I(M7)	Vm(7),I(7)	IV(7)	♯IV◦7	I(M7)	VI(7)	IIm(7)	V(7)	I(M7)	I(M7)
c)	I(M7)	IV(7)	I(M7)	Vm(7),I(7)	IV(M7)	IVm(7)	IIIIm(7)	VI(7)	IIm(7)	V(7)	I(M7)	I(M7)
d)	I(M7)	IIm(7),♯II◦7	IIIIm(7)	Vm(7),I(7)	IV(M7)	IVm(7),bVII(7)	IIIIm(7)	bIIIIm(7)	IIm(7)	V(7)	I(M7)	I(M7)
e)	I(M7)	VII(♭7),III(7)	VIm(7),II(7)	Vm(7),I(7)	IV(M7)	IVm(7),bVII(7)	bIII(M7)	bIIIIm(7),bVI(7)	IIm(7)	V(7)	I(M7)	I(M7)
f)	I(M7)	IV(7)	I(M7)	bIIm(7),bV(7)	IV(7)	♯IV◦7	IIIIm(7)	VI(7)	IIm(7), V(7)	bVIm(7),bII(7)	I(M7)	I(M7)
g)	bII(7),bV(7)	VII(7),III(7)	VI(7),II(7)	V(7),I(7)	IV(7)	♯IV◦7	IIIIm(7)	bIII(7)	VIm(7)	bII(7)	I(M7)	I(M7)

Figure 1: Some Jazz 12-bars (adapted from Coker, 1964)

“Phonological Spelling” of Chords

- We can regard all of the chords in figure 1 as falling into four basic chord types within which they differ only as to which particular additional notes they add. The four types are distinguished as to whether they are the major chord X or the minor chord X_m , and whether they include the dominant seventh note, when they are written X^7 and X_m^7 respectively

$$\begin{aligned} (1) \quad a. \quad & \{X(M7), X(7), X(9), X(13)\} & := & X \\ & b. \quad \{X_m(7), X_m(6)\} & := & X_m \\ & c. \quad \{X(7), X(b9), X(b10), X(7+5)\} & := & X^7 \\ & d. \quad \{X_m(7), X_m(9), X(\phi 7)\} & := & X_m^7 \end{aligned}$$

- The “dominant seventh” chords X^7 and X_m^7 create an expectation of IV_X the chord of the 4th degree of the scale relative to X as I or tonic. X and X_m do not create a particular expectation of this kind.

Phonological Spelling of Chords (Contd.)

- A dominant seventh chord followed by that expected IV chord is an elementary perfect or authentic **cadence**.
- The phonological spelling level of rules is where issues of voice-leading and inversion should be brought into the grammar, analogously to processes like liaison and lenition in speech.
- It is also where issues of ambiguity come in— X and X^7 can both be realized as $X(7)$, since the minor seventh and dominant seventh are homophones in equal temperament.

The Recursive Nature of Cadences

- You can derive more complicated blues chord sequences from simpler ones by propagating authentic cadences backwards. That is, successive substitutions in the basic skeleton (a) of Figure 1 generate examples like those in Figure 2, in which the elaborated cadence is underlined:

<i>a.</i>	I	IV	I	I ⁷	IV	IV	I	I	V ⁷	V ⁷	I	I
<i>a'</i>	I	IV	I	I ⁷	IV	IV	I	I	IIIm ⁷	V ⁷	I	I
<i>a''</i>	I	IV	I	I ⁷	IV	IV	I	VI ⁷	IIIm ⁷	V ⁷	I	I
<i>a'''</i>	I	IV	I	I ⁷	IV	IV	IIIIm ⁷	VI ⁷	IIIm ⁷	V ⁷	I	I
..... <i>etc.</i>												

Figure 2: Recursive Propagation of the Authentic Cadence

Cadences and Cascaded Finite-State Transducers

- The value of, for example, the $\text{III}m^7$ chord in a''' in Figure 2 is therefore dependent upon a chain of substitutions working back from a quite distant V^7 to its right, suggesting a *right-branching* tree-structure characteristic of many Schenkerian approaches and a right-to-left parsing algorithm, like that in Winograd 1968.
- Steedman 1984 (hereafter *GGJ*) proposes a set of substitution rules, which are most naturally interpreted as the rules of a finite-state transducer (FST), which can be used in cascade in a left-to-right pass to restore the chords of the I, IV, V 12-bar skeleton. (Cf. Pachet 2000 and Chemillier 2004).
- However, we want to know more than that ‘Solar’ is (or is not) a Blues. We want to know *why* it is (or isn’t).
- That is, we want a **harmonic analysis**, an important component of what a musical piece *means*.

Longuet-Higgins' Theory of Tonal Harmony

- To make the grammar deliver a harmonic analysis, we need to get away from the whole idea of *syntactic* substitution of one chord for another that underlies *GGJ* and Johnson-Laird 1991, and to seek a grammar founded more straightforwardly in a musical semantics or “model theory” for harmony.
- The key to achieving this lies in work by Longuet-Higgins 1962a, 1962b, who showed that the harmonic relation between a pair of notes can be expressed as a vector in a three-dimensional discrete space whose generators are respectively related to integer frequency ratios of two, three, and five, and no others.
- It will be convenient to project this three dimensional space onto two dimensions along the “times two” axis, since this corresponds to the octave.

E	B	F#	C#	G#	D#	A#	E#	B#
C	G	D	A	E	B	F#	C#	G#
Ab	Eb	Bb	F	C	G	D	A	E
Fb	Cb	Gb	Db	Ab	Eb	Bb	F	C
Dbb	Abb	Ebb	Bbb	Fb	Cb	Gb	Db	Ab

Figure 3: (Part of) The Space of Note-names (Longuet-Higgins 1962a,b)

III^-	VII^-	$\#\text{IV}^-$	$\#\text{I}$	$\#\text{V}$	$\#\text{II}$	$\#\text{VI}$	$\#\text{III}^+$	$\#\text{VII}^+$
I^-	V^-	II^-	VI	III	VII	$\#\text{IV}$	$\#\text{I}^+$	$\#\text{V}^+$
bVI^-	bIII^-	bVII^-	IV	I	V	II	VI^+	III^+
bIV^-	bI^-	bV^-	bII	bVI	bIII	bVII	IV^+	I^+
bbII^-	bbVI^-	bbIII^-	bbVII	bIV	bI	bV	bII^+	bVI^+

Figure 4: (Part of) The Space of Disambiguated Harmonic Intervals

Harmony Theory (Notes)

- In this figure the intervals are disambiguated. The prefix \sharp and \flat roughly correspond respectively to the traditional notions of “augmented” intervals, and to “minor” and/or “diminished” intervals, while the superscripts plus and minus roughly correspond to the “imperfect” intervals.
- This is a musical true fact, which is obscured by modern equally tempered tuning, which equates all positions separated by the Comma of Didymus (yielding the spiral space of standard notation explored by Chew 2001), as well as those related by an augmented seventh (yielding the toroidal space of equal temperament).

Chord Progression

- Musically coherent chord sequences such as the twelve-bar blues have something to do with orderly progression to a destination by small steps in this space.
- For example, the basic sequence in Figure 1a, repeated as Figure 2a, is a closed journey around a central I visiting the immediately neighbouring IV and V. Figure 2a' makes a jump to the right to II, then returns via V.
- The work of moving the harmonic reference point around in the space is mainly done by **dominant seventh** chords.

	(II ⁻)		III	(VII)		
	bVII ⁻	(IV)	I	(V)	(II)	
						(IV ⁺)
	(bbIII ⁻)		(bIV)	(bI)		

Figure 5: The Dominant Seventh Chord (circles) and its resolution

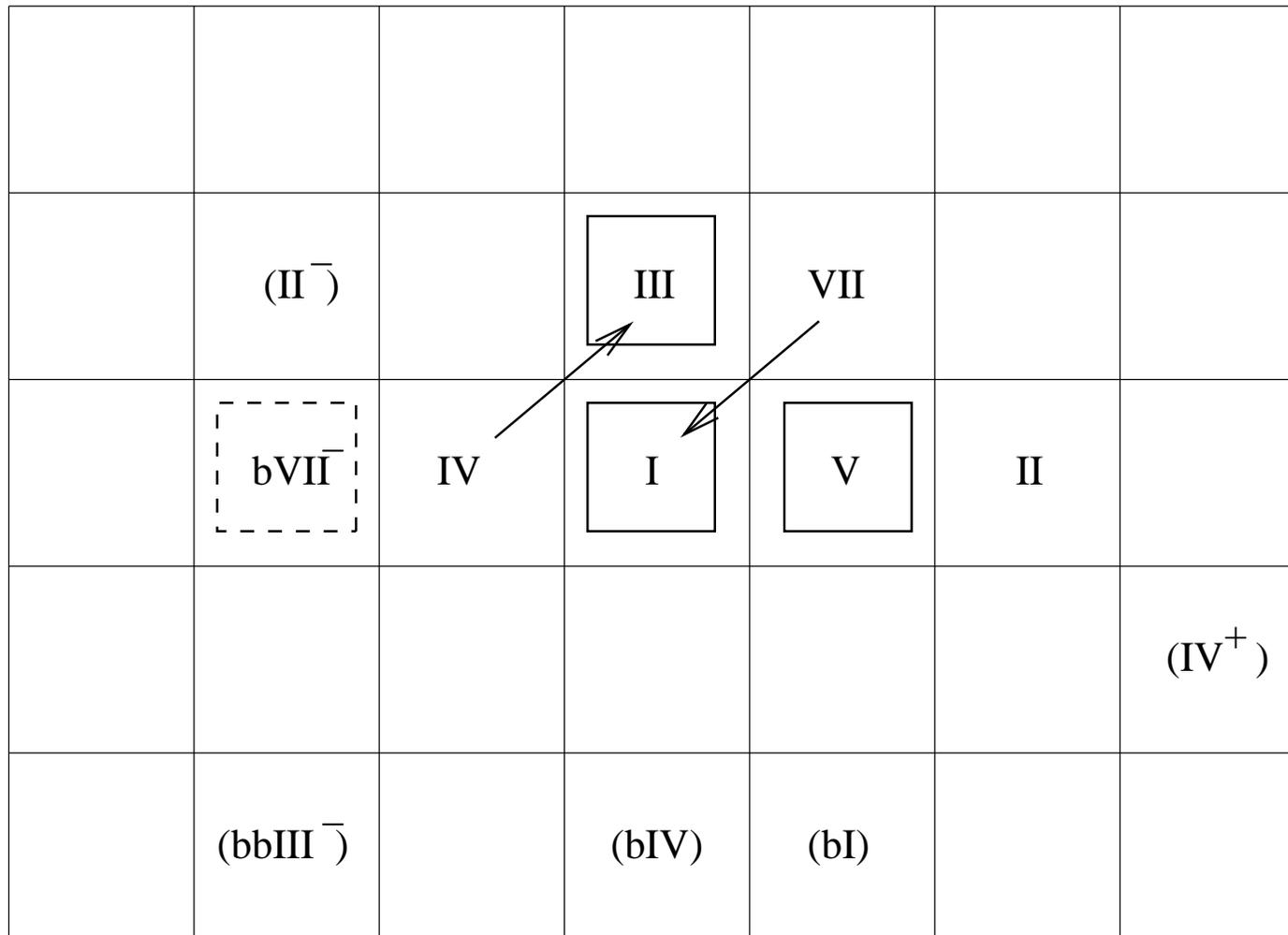


Figure 6: The Dominant Seventh Chord and its resolution (squares)

The Dominant Seventh

- The representation makes it obvious why the harmonically closest interpretations of the IV and the II are not any of the imperfect or diminished alternatives shown in brackets. It is the addition of the dominant seventh of V, the circled IV, that makes the V chord have a hole in its middle, into which a triad on I (squared I, III, and V) fits neatly, sharing one note with the first chord, and with the two remaining notes standing in semitone “leading note” relations. (There are voice-leading implications here.)
- This is a different kind of disambiguation, of individuals in the model, comparable to resolution of pronoun reference in natural language.

Another Chord that Works like a Dominant Seventh

- If we substitute a $\flat\text{II}m^7$ for the $\text{V}7$ in the chord sequence of Figure 2, to obtain a final cadence $\text{III}m^7 / \text{VI}^7 / \text{II}m^7 / \flat\text{II}m^7 / \text{I} / \text{I}$, the effect of a recursive perfect cadence seems unaffected.^a
- This is because it is not really a $\flat\text{II}$ chord, but a V chord in which the $\flat\text{II}$ note is heard as the tritone of V . The supposed minor mediant of $\flat\text{II}$, $\flat\text{IV}$ is heard as III , the mediant or major seventh of the destination I , while the supposed minor seventh of $\flat\text{II}$, $\flat\text{I}$ is heard as VII , the mediant of V .
- The reason that this function is not made explicit by notating it as $\text{V}\flat 2, 6, \flat 5$ is to do with the fact that the root V is actually omitted for reasons of consonance and voice-leading. ($\flat\text{II}m^7$ is also a lot easier for the performer to take in in real time.)
- We can again see why it is disambiguated in this way by the succeeding I by looking at the progression in the harmonic space:

^aThanks to Richard Terrat for discussions on this question.

			III	VII		
			I	V		
(bI ⁻)		bII	bVI			
			(bIV)	(bI)		

Figure 7: The Dominant Tritone Chord (circles) and its resolution

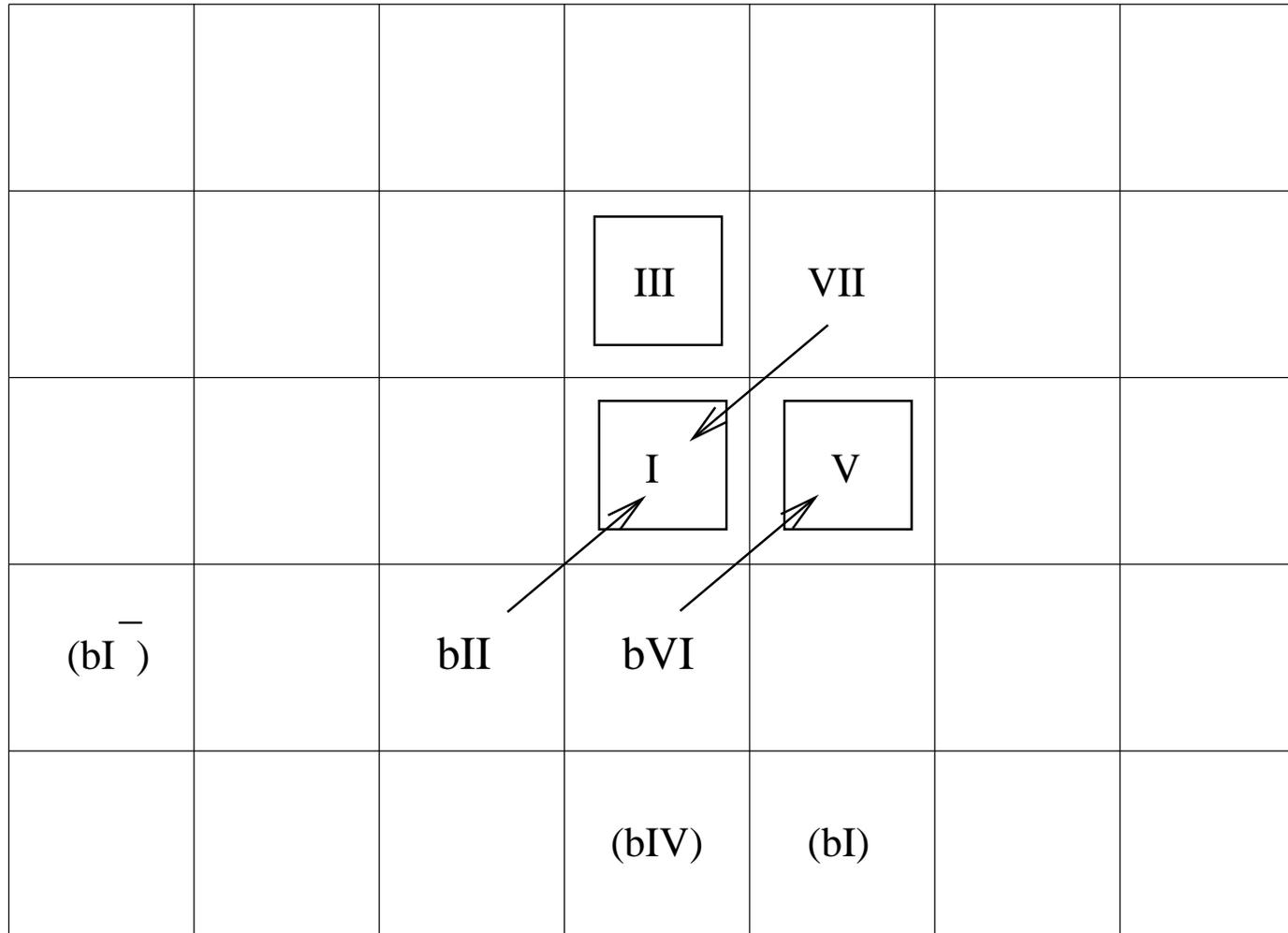


Figure 8: The Dominant Tritone Chord and its resolution (squares)

Combinatory Categorical Grammar for English

- Now we need a syntax that is capable of supporting the musical semantics for use in a parser. We can borrow it from natural language.

$$\begin{aligned} (2) \quad & \cancel{S} \quad \rightarrow \quad \cancel{NP} \quad \cancel{VP} \\ & \cancel{VP} \quad \rightarrow \quad \cancel{TV} \quad \cancel{NP} \\ & \cancel{TV} \quad \rightarrow \quad \{ \cancel{\text{eats}}, \text{drinks}, \dots \} \end{aligned}$$

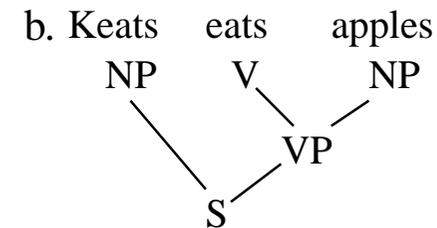
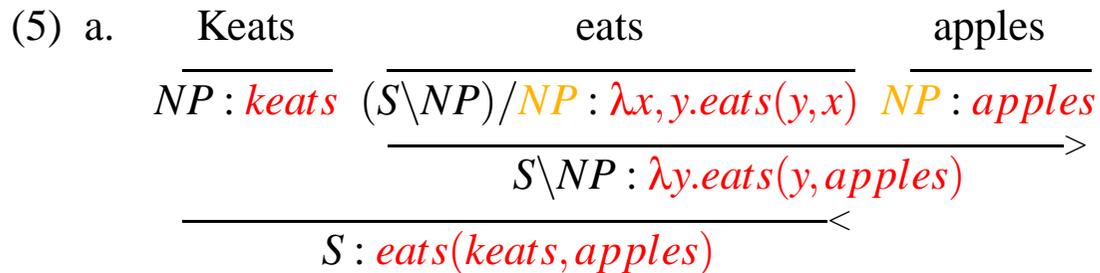
$$(3) \quad \text{eats} := (S \backslash NP) / NP$$

(4) *Functional Application:*

$$a. \quad X/Y \quad Y \quad \Rightarrow \quad X$$

$$b. \quad Y \quad X \backslash Y \quad \Rightarrow \quad X$$

CCG for English



- (The annotations $>$ and $<$ on combinations in a, above, are mnemonic for the rightward and leftward function application rules 4a,b).
- In order to capture linguistic phenomena such as coordination, relativization, intonation structure, and word order in languages other than English, CCG adds syntactic operators related to the Combinators of Schönfinkel and Curry

Composition and Type-raising

(6) *Forward Composition* ($> \mathbf{B}$):

$$X/Y \quad Y/Z \quad \Rightarrow_{\mathbf{B}} \quad X/Z$$

- Together with “type-raised” categories that can be substituted in the lexicon or introduced by rule for argument categories like subject NPs, rule 6 allows extractions as follows:

$$\begin{array}{ccccccc}
 (7) & \text{(a man)} & & \text{who(m)} & & \text{I} & & \text{like} \\
 & \underline{\hspace{2cm}} & & \underline{\hspace{2cm}} & & \underline{\hspace{2cm}} & & \underline{\hspace{2cm}} \\
 & & & (N \setminus N) / (S / NP) & & S / (S \setminus NP) & & (S \setminus NP) / NP \\
 & & & & & \underline{\hspace{2cm}} & & \underline{\hspace{2cm}} > \mathbf{B} \\
 & & & & & S / NP & & \\
 & & & \underline{\hspace{6cm}} & & \underline{\hspace{6cm}} & & \underline{\hspace{6cm}} > \\
 & & & N \setminus N & & & &
 \end{array}$$

Substitution

(8) *Forward Substitution* ($> \mathbf{S}$):

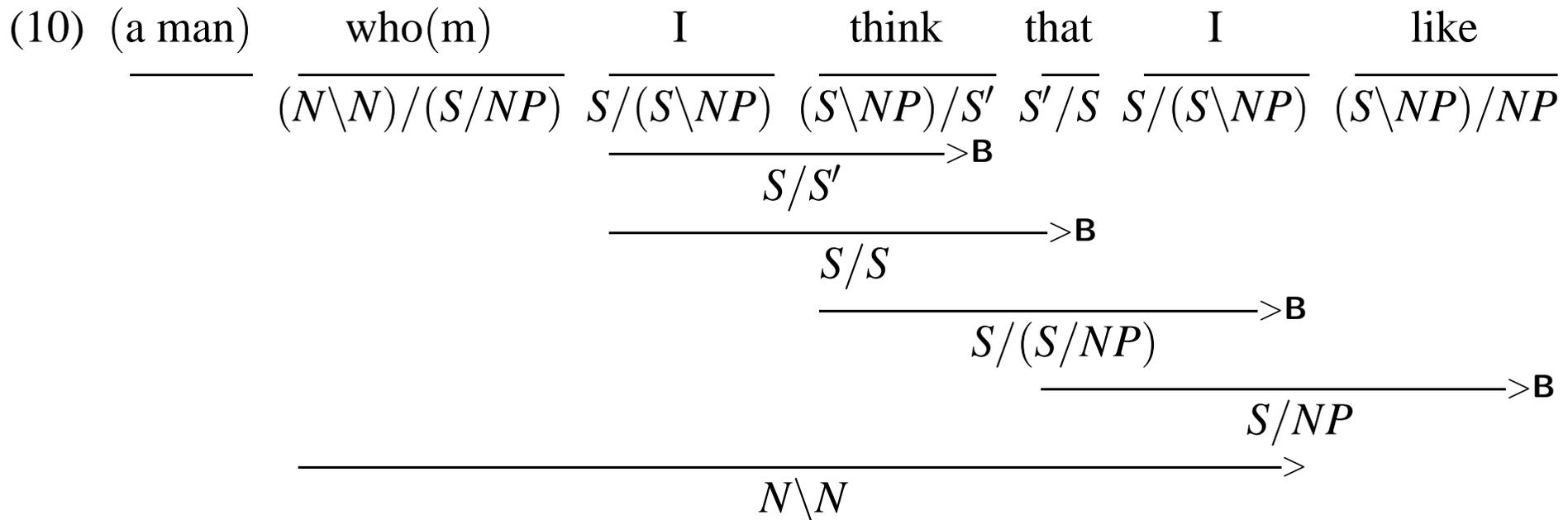
$$(X/Y)/Z \quad Y/Z \quad \Rightarrow_{\mathbf{S}} \quad X/Z$$

- Rule 8 allows extractions like the following:

$$\begin{array}{ccccccc}
 (9) & \text{(a man)} & & \text{who(m)} & & \text{Every acquaintance of} & & \text{dislikes} \\
 & \underline{\hspace{2cm}} & & \underline{\hspace{2cm}} & & \underline{\hspace{2cm}} & & \underline{\hspace{2cm}} \\
 & (N \setminus N) / (S / NP) & & (S / (S \setminus NP)) / NP & & (S \setminus NP) / NP & & \\
 & & & \underline{\hspace{10cm}} & & & & \text{> } \mathbf{S} \\
 & & & S / NP & & & & \\
 & \underline{\hspace{15cm}} & & & & & & \text{>} \\
 & & & N \setminus N & & & &
 \end{array}$$

Composition and Type-raising Contd.

- Such extractions are immediately predicted to be unbounded:



- CCG also predicts that fragments like *I like* and *I think that I like* are able to undergo coordination and be made intonational phrases.

Properties of Combinatory Categorical Grammar

- While we will not go into the details here, it is also crucial that the combinatory rules allow the immediate assembly of a correct semantic interpretation for unboundedly extending non-standard constituents like *I think that I like* (Steedman 2000).
- This is not surprising given the resemblance to Combinatory Logic.
- Indeed, the low expressive power of CCG (which is only “mildly” context sensitive—see Joshi et al. 1991) come from syntactic constraints on permissible rules—see Steedman 2000.
- The interesting thing about such grammars for present purposes is that they allow left-branching analysis of structures like the English clause, which we usually think of as predominantly right-branching.
- There are implications for incremental interpretation in language and music under the “Strict Competence” Hypothesis.

A Categorical Chord Grammar

- 1a. $X := I_X$
- 1b. $Xm := I_Xm$
- 2a. $X := V_X \setminus V_X$
- 2b. $Xm := V_Xm \setminus V_Xm$
- 3a. $Xm^7 := I_Xm^7 / IV_X^7$
- 3b. $X^7 := I_X^7 / IV_X(m)^7$
- 4. $Xm^7 := \#IV_Xm^7 / VII_Xm^7$
- 5. $Xm := \flat VII_Xm \setminus \flat VII_Xm$
- 6. $X \circ 7 := \begin{array}{l} \flat V_X / \flat V_X \\ \flat II_X / \flat II_X \\ \flat VII_Xm^7 / \flat VII_Xm^7 \end{array}$

Figure 9: A Categorical Chord Grammar

- I_X, V_X , etc. are I, V etc. relative to the root X on the left of the $:=$. Brackets round the minor (m) mean that if the basic chord is minor then so is the categorial type.
- Most categories are simply the identity function, but 3a and 3b do the real work of elaborating the perfect cadence. The reason they are I^7/IV^7 rather than I^7/IV is to do with recognizing the end of the cadence—see below.
- Rule 4, defining a Xm^7 as a dominant tritone chord is also non-trivial.
- We add a pair of trivial syncategorematic rules resembling coordination that make sequences of X s into a single X . These have the property of passing the 7 marker to the rightmost daughter (brackets here mean the $(^7)$ is optional).

$$(11) \quad X(m) \quad X(m) \quad \Rightarrow \quad X(m) \quad (\langle \& \rangle)$$

$$X(m)(^7) \quad X(m)^7 \quad \Rightarrow \quad X(m)^7 \quad (\langle \& \rangle)$$

- This notation can easily be augmented to enforce the condition that the combination of an X/Y occupying a bars with a Y occupying b bars yields an X occupying $a + b$ bars. This detail is omitted.

The Categorical Grammar of Authentic Cadences (Contd.)

- Unlike GGJ, this fragment does not work by substitution on a previously prepared skeleton.
- It is still incomplete, in that it does not yet specify the higher levels of analysis that stitch the sequences of cadences together into canonical forms like twelve-bars, and variations on *I Got Rhythm*.
- Stochastic POS tagging techniques are likely to do very well at disambiguating homophones like $X(7)$ chords. ($X(7)$ is likely to be X^7 if followed by IV_X , X if followed by V_X , etc.)
- Just as in the linguistic grammar, we can associate a semantic interpretation with categories, which the combinatory rules will “project” onto derivational structure, as follows:

The Categorical Grammar with Semantics

- 1a. $X := I_X : X$
- 1b. $Xm := I_Xm : X$
- 2a. $X := V_X \setminus V_X : \lambda x.x$
- 2b. $Xm := V_Xm \setminus V_Xm : \lambda x.x$
- 3a. $Xm^7 := I_Xm^7 / IV_X^7 : \lambda x.leftonto(x)$
- 3b. $X^7 := I_X^7 / IV_X(m)^7 : \lambda x.leftonto(x)$
4. $Xm^7 := \#IV_Xm^7 / VII_Xm^7 : \lambda x.leftonto(x)$
5. $Xm := \flat VII_Xm \setminus \flat VII_Xm : \lambda x.x$
6. $X \circ 7 := \flat V_X / \flat V_X : \lambda x.x$
 $\flat II_X / \flat II_X : \lambda x.x$
 $\flat VII_Xm^7 / \flat VII_Xm^7 : \lambda x.x$

Figure 10: The Categorical Grammar with Semantics

The Categorical Grammar with Semantics

- Just as in the linguistic grammar, we can associate a semantic interpretation with categories, which the combinatory rules will “project” onto derivational structure.
- The rules that make sequences of X s into a single X are also identity functions of a sort:

$$(13) \quad X(m) : x \quad X(m) : x \quad \Rightarrow \quad X(m) : x \quad (\langle \& \rangle)$$

$$(14) \quad X(m)^{(7)} : x \quad X(m)^7 : x \quad \Rightarrow \quad X(m)^7 : x \quad (\langle \& \rangle)$$

A Derivation (Contd.)

- A cadence requires an origin as well as a destination.
- Instead of just applying an extended cadence to its target, we will give the target a higher-order type that labels the result explicitly as $I_X \setminus I_X$, the category of a non-initial cadential modifier of I_X , via the following rule reminiscent of type-raising:

$$(17) \quad 7. \quad X \rightarrow (I_X \setminus I_X) \setminus (Y^7 / I_X^7)$$

- Semantically we can think of the rule as follows

$$(18) \quad 7. \quad X : \textit{origin} \rightarrow (I_X \setminus I_X) \setminus (Y^7 / X^7) : \lambda \textit{cadence} . \lambda \textit{origin} . \textit{origin} + \textit{cadence}(\textit{origin})$$

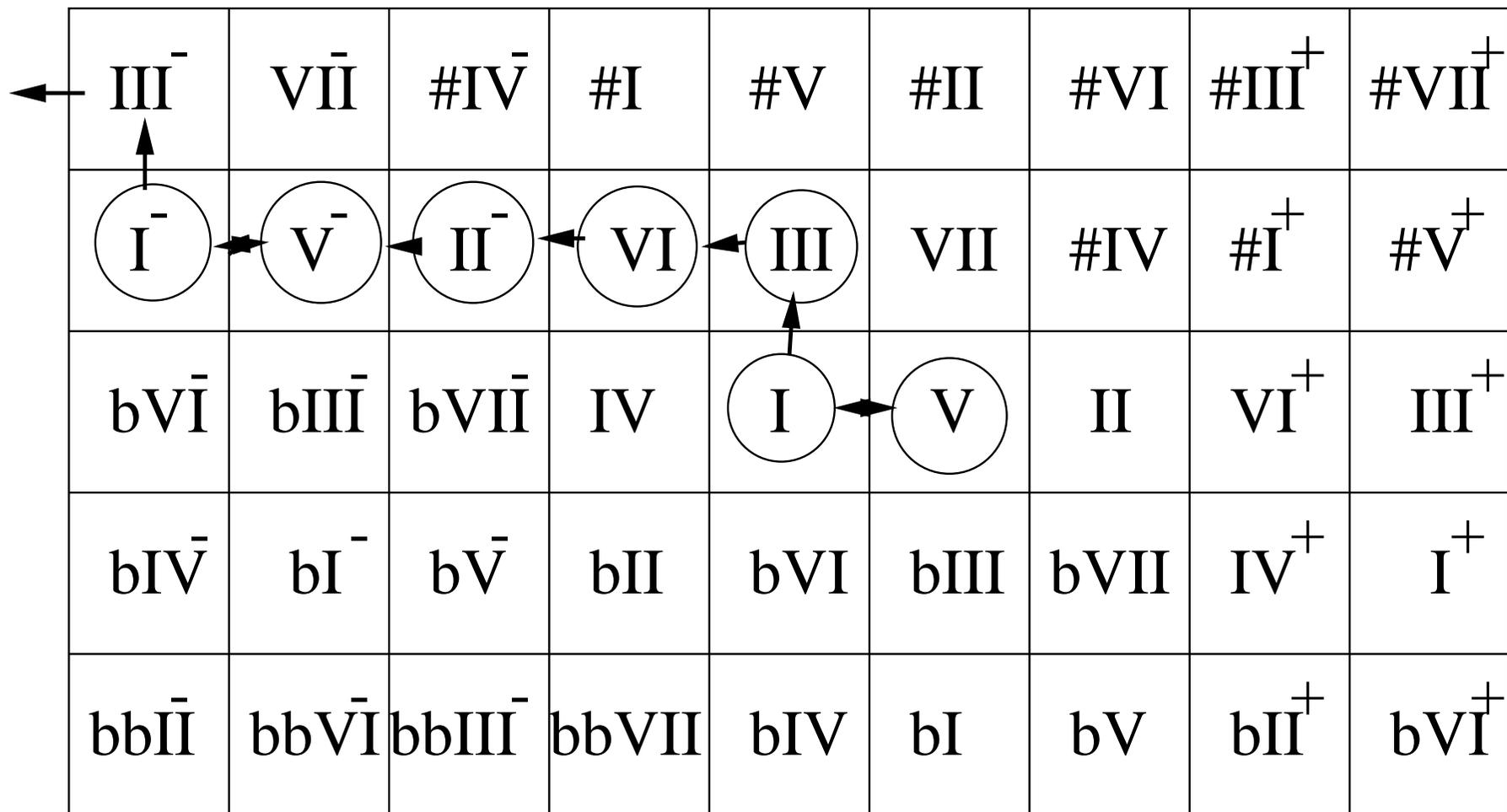


Figure 11: The Denotation of an extended Authentic Cadence (Basin St. Blues)

A Model for the Interpretation

- This denotation, which corresponds to Figure $2a'''$, is interesting, because it takes a step up to III , then proceeds via leftward steps to end up on I^- .
- I^- is musically distinct from the original I , and if perfectly intoned (as opposed to being played on an equally tempered keyboard), would differ from the original in a ratio of 80:81.
- So it is only practicable to play this sort of music in Equal Temperament.
- When you do so, the new I sounds **sharp**. (There is a feeling of elation.)

Extending the Grammar to the Plagal Cadence

- It is a prediction of the theory that the plagal ($IV-I$) cadence will be elaborated in a similar way, to give sequences which are the mirror image of the authentic cadence, with a “right-onto” semantics.
- We can do this by introducing the following two categories parallel to 3a and 3b:

$$3'a. \quad X_m \quad := \quad I_X m / V_X : \lambda x. \text{rightonto}(x)$$

$$3'b. \quad X \quad := \quad I_X / IV_X(m) : \lambda x. \text{rightonto}(x)$$

Figure 12: The Plagal Cadence Categories

- While the plagal cadence is less commonly exploited than the authentic, this prediction is correct: “Hey Joe” (Hendrix, 1965) is an exact plagal spatial mirror image of figure 11.
- In equally tempered tuning, the new tonic sounds **flat**, and depressing.

III^-	VII^-	$\#\text{IV}^-$	$\#\text{I}$	$\#\text{V}$	$\#\text{II}$	$\#\text{VI}$	$\#\text{III}^+$	$\#\text{VII}^+$
I^-	V^-	II^-	VI	III	VII	$\#\text{IV}$	$\#\text{I}^+$	$\#\text{V}^+$
bVI^-	bIII^-	bVII^-	IV	I	V	II	VI^+	III^+
bIV^-	bI^-	bV^-	bII	bVI	bIII	bVII	IV^+	I^+
bbII^-	bbVI^-	bbIII^-	bbVII	bIV	bI	bV	bII^+	bVI^+

Figure 13: The Denotation of an Extended Plagal Cadence (*Hey Joe*)

Conclusion

- The grammar now interprets twelve-bar and other sequences as being made up of cadences.
- An important result from the point of view of psychological plausibility is that the derivation that produces this very orthodox right branching cadential semantics is predominately *left* branching.
- To that extent, it is also semantically *incremental*, delivering an interpretable result at each reduction, more or less chord by chord.
- This property is likely to be important form musical “language modeling” for category disambiguation in this very ambiguous musical language.
- Is this “mildly context-sensitive” expressive power necessary?

A Musical Parasitic Gap?

- N. Washington/B. Kaper (1947) *On Green Dolphin Street*

[When I recall the love I found on]_{(S/S)/NP}

[Dm7 Bm7b5 E7b9 Am7]_{(Dm7/F#m7)/Dm}

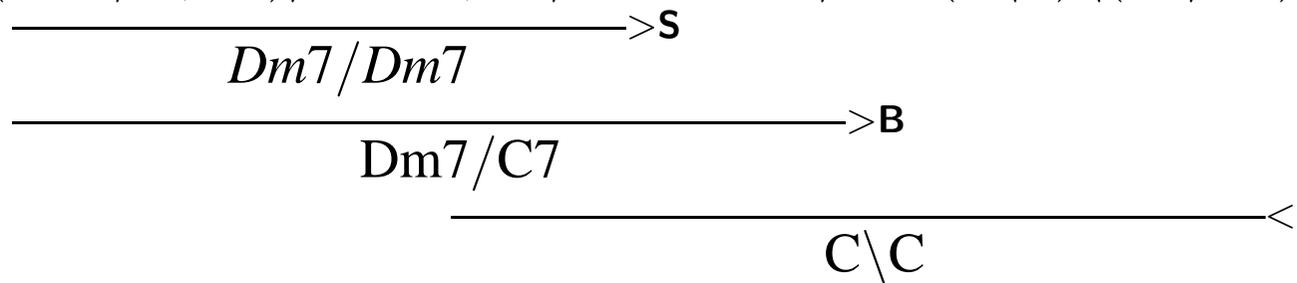
[I could kiss the ground on]_{S/NP}

[F#m7 B7 Em7 A7]_{F#m7/Dm7}

[Green Dolphin Street]_{S\ (S/NP)}

[Dm7 G7]_{Dm7/C7} C_{(C\C)\ (Y7/C7)}

(21) $(Dm7/F\#m7)/Dm$ $F\#m7/Dm7$ $Dm7/C7$ $(C\C)\ (Y7/C7)$



- Are there musical crossed dependencies?

References

Chemillier, Marc, 2004, Grammaires, automates, et musique, in Briot, J.-P. and Pachet, Francois, (eds.) *Informatique musicale*, Hermès, Paris

Cohn, Richard, 1998, Introduction to Neo-Riemannian Theory: A Survey and a Historical Perspective, *Journal of Music Theory*, **42**, 167-180,

Ellis, A. 1874 On Musical Duodenes. *Proceedings of the Royal Society*, **23**,

Johnson-Laird, P.N. 1991 Jazz Improvisation: a Theory at the Computational Level, in Howell *et al.*, 1991, p.291-326.

Longuet-Higgins, H.C. 1962a, Letter to a Musical Friend. *The Music Review*, **23**, 244-248.

Longuet-Higgins, H.C. 1962b, Second Letter to a Musical Friend. *The Music Review*, **23**, 271-280.

Lovelace, Ada Countess, 1842, “Translator’s notes to an article on Babbage’s Analytical Engine,” in R. Taylor (ed.) *Scientific Memoirs*, 3. 691-731.

Pachet, Francois, 2000, Computer Analysis of Jazz Chord Sequences: Is Solar a Blues? , in Eduardo Miranda, (ed.) *Readings in Music and Artificial Intelligence*, Harwood Academic Publishers, New York

Steedman, M. 1984 A Generative Grammar for Jazz Chord Sequences. *Music Perception*, **2**, 52-77.

Steedman, M. 1996, The Blues and the Abstract Truth: Music and Mental Models, in J. Oakhill and A. Garnham, (eds.), *Mental Models in Cognitive Science*, Erlbaum. 305-318.

Steedman, M. 2000 *The Syntactic Process*, MIT Press

Winograd, T. 1968 Linguistics and the Computer Analysis of Tonal Harmony, *Journal of Music Theory*, **12**, 2-49.