Move in Syntax, --logically necessary or undefinable in the best case?

1. Move as Logical Necessity

Consider the argument in (1):*

(1) Given 'merge' of x, y, x and y either overlap or not. Hence it is logically necessary that both these situations must be individuated in our theory of syntax. Call the former "external merge" the latter "internal merge" or 'move'. Hence 'move' must exist (in this sense).[1]

In other words given (syntactic) 'merge', (syntactic) 'move' is logically necessary. In what follows I would like to look more closely at this claim. The wording of (1) might seem strange: how could an empirical fact follow from a logical necessity? I assume that the argument is in fact based on two empirical assumptions: merge exists and the term '(non-)overlap' is part of the (minimalist) syntactic vocabulary.

I shall omit single quotes around "merge" and "move" below, but I adopt these terms only for the sake of discussion. I do not think that competence theory is derivational (cf. Brody 1995, 1997, 1998, 2002 etc.). This issue is orthogonal however to the present discussion.

2. Self-attachment, Sideways Movement

If the merged categories x, y overlap then either (a) they are identical or (b) one is identical to some proper part of the other or (c) there is a category that is a proper part of both x and y. (Note that (b') where both x and y are proper parts of the other cannot exist if/since infinite regress is not allowed.)

The second case, (b), is the typical case of internal merge. (a) would correspond to "self-attachment" (see e.g. Bury 1993 for some discussion) and (c) to sideways movement (Nunes 2001, 2004) The status of such configurations is moot, --in particular whether they should fall under internal merge is not fully clear. This matter is also largely tangential to my concerns and I shall leave it open here.

3. Is "Non-overlap" in the minimalist vocabulary?

Consider then, if we have independent reasons to think that "(non-)overlap" is in the minimalist, conceptually (near-)necessary, vocabulary, i.e. that it is a logically or conceptually necessary notion for syntactic theory. One obvious way of thinking about (non-)overlap would be to compose it (as above in section 2.) from the identity and the part-whole relations. Syntax clearly needs some notion of part-whole relation, in a good sense that is what syntax is about. Identity also appears to be an apparently inevitably necessary notion. While the claim that composing these two relations is a logical necessity may well be questioned, this seems only a rather minor and probably inconsequential objection to (1). It would appear, that composing these two necessary relations into the concept of (non-)overlap comes at no or minimal cost.

On a slightly closer examination, however, the necessity of the relevant notion of identity becomes less clear. Presumably, some distinction between numerical, token identity and lack of it is indeed necessary for any complex structure to exist. Type identity is different however and is not relevant to syntax in any obvious or necessary way. In fact this appears to be an interpretive concept. We know that type identity is necessary in the interpretive component, quite independently of issues connected with '(internal) merge'. This is obvious if the syntax-interpretation interface contains chains whose members are copies. But even apart from chains, that contain a single argument for thematic interpretation, type identity appears to be necessary for the interpretation of relations that link multiple arguments like binding and control phenomena or potentially link multiple contentive elements like agreement.

Note, incidentally, the potential argument for LF chains here: the notion of type identity, hence the more abstract chain-theoretical notion of argument (or contentive) apparently comes for free in interpretation. But even more relevantly for present purposes, the fact that type identity is necessary in the interpretive component does not necessarily mean under a modular approach that it is available also in syntax. Arguments would be necessary to import (some aspects of?) type identity (of copies) into narrow syntactic theory. I do not think that strong arguments have ever been provided for this or even that the issue has ever been properly addressed. Furthermore, the direction strikes me as dubious (particularly in a minimalist type setting) given the apparently interpretive nature of the type-identity relation and the duplication that such a move might involve (Brody 1998, 2000b, 2001, 2002). But whatever the status of the arguments for syntactic type-identity, I cannot see that a reference to the logical status of (non-)overlap between the elements merge operates on (i.e. (1) above), could be among them. It seems clear that type-identity is not a logically necessary syntactic notion. Hence neither is the notion of chain (or move) if this is based on the concept of type-identity.

4.Multiple Domination

On the other hand the McCawley-Hudson-Starke theory of multiple domination (re-merge) makes use of numerical identity between the copies of 'internal merge', hence the previous considerations relating to the non-necessity of the concept of type-identity in syntactic theory would not carry over unchanged. On the multiple domination approach the components of the nondistinctness relation (token identity and part-whole relation) indeed seem conceptually necessary. Clearly, however, the multiple domination approach itself is not a pure logical necessity, the assumption that a category may be re-merged (immediately dominated by more than one category) is a factual one, in principle it could be true or false. But perhaps the possibility of re-merge is indeed the default minimal assumption? Have we then found a potentially valid reconstruction of the claim in (1) in the context of the re-merge theory?

(1') Given 'merge' of x, y, x and y either overlap or not. Hence it is the minimal assumption that both of these situations are individuated in our theory of syntax. Call the former "external merge" the latter "re-merge" or 'move'. Hence in the best case 'move' must exist (in this sense).

This line of thought appears to lead into various empirical problems. Reconstruction phenomena provides direct evidence against re-merge: for example lack of reconstruction effects in adjuncts appears to show that there is no strict identity between the moved element and its trace contrary to what the numerical identity of the re-merge theory would predict. Importantly, the constituent that appears to be missing from the trace position is not always peripheral. Hence no constituency respecting analysis can resolve the problem by claiming that numerical identity holds between the trace and moved element with the missing constituent added only in the moved position. For example in (2), given the well motivated assumption that principle C is an "everywhere condition", "Martin" must not be present in the trace position to avoid the principle C violation.

(2) Which claim that the woman that Martin/x telephoned met Mary did you think he/x denied

But neither "Martin" nor the adjunct containing it can apparently be added late without violating either cyclicity or constituency. Notice that if we do not adopt the multiple domination approach then we have at least two occurences of "Martin" in (2), which need not be fully identical. The occurence c-commanded by the coreferential element can therefore be a pronoun, 'nondistinct' from the R-expression (Brody 1995).

(I do not intend to claim here that multiple domination must be observationally inadequate. To achieve observational adequacy, as a reviewer points out, one could for example exempt R-expressions that are internal to multiply dominated adjuncts from the requirement that principle C must be obeyed with respect to all attachments, that is from the "everywhere" nature of principle C. This clearly does not appear to be a tempting direction to pursue.)

Furthermore there are various configurations of agreement and control in addition to move-chain that can be treated in terms of type-identity. If we continue to analyse these phenomena in terms of type identity, then the minimal hypothesis of (1') leads to a system where we have two ways of resolving what on the relevant level of abstraction appears to be a unitary phenomenon. We will then need to use both multiple dependencies of a single element and type identity. While there is nothing logically impossible about this, it is not clear why language should make use of two different solutions to express identity. If we say it does, this would raise the question of whether the thinking in (1') indeed leads to a minimal system. On the other hand treating control in terms of numerical identity, would mean treating control on a par with move, and would therefore inherit all the numerous problems of such an approach (see Brody 2000b, 2001, Hornstein 1998, 2000, Landau 2003, Boeckx and Hornstein 2003 for some discussion and opposing views). Treating all agreement phenomena in terms of numerical identity is likely to create many further difficult issues.

It is clear, however, that if the existence of multiple domination indeed represents the minimal hypothesis then the most interesting direction is this apparently difficult one: we have to face the empirical problems and hope to resolve them.

5. Transitive Membership as a Primitive

I would like to suggest instead that the existence of re-merge is the minimal hypothesis only in the context of a theory of categorial ("phrase") structure that itself is not minimal. The theory of categorial structure that re-merge presupposes is in fact stronger (more expressive) than necessary. Once we properly restrict (the expressive power of) our theory of structure, the existence of multiple domination is not going to be the minimal hypothesis, multiple domination will be inexpressible. This will entail that the only construction involving identity language makes crucial use of is the type-identity relation in the interpretive component. The empirical problems created by numerical identity theories of move (and of control and agreement) will not arise.

Thinking of syntactic trees as set-theoretical objects, we normally take a category C to be a set whose members are categories dominated by C. Intuitively, the notion "member of x" may be taken in a transitive or intransitive sense, where standard set theory opts strictly for the latter. Let me refer to the former sense as "t-member" and the latter, standard sense "it-member". I suggest to experiment with the former interpretation, ie. where we assume the transitivity axiom in (3)

(3) if X t-member of Y and Y t-member of Z then X t-member of Z.

Let us assume also the identity axiom in (4), --where the clause "if there is a Q st. Q is t-member of V" is used to make the condition cover only non-atomic categories.

(4)if there is a Q st. Q is t-member of V,
then V=W iff for all Z, if Z t-member of V then Z t-member of W

The transitivity and the identity conditions restrict the interpretation of the notion of t-member and serve as axioms on the basis of which we can individuate any nonatomic element. Notice that I am not defining anything in terms of standard set-theory but rather suggesting a different interpretation of the notion of set-membership.

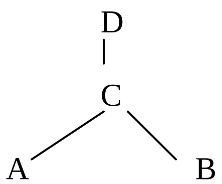
To make this more intuitive, suppose we have containers and things in containers. Some of these things in containers may themselves be containers. Suppose in a big container we have a blue marble and another smaller container that contains a white marble. The natural language question: "what does the big container contain" can be truthfully answered quite naturally in at least three different ways. We can say it contains two marbles, call this the m(ereological)-answer. Or we can say that it contains a blue marble and a smaller container, the non-t answer nearest in spirit to standard set theory. Finally we could also say that it contains a blue marble, a smaller container and a white marble (which is contained also in the smaller container), -the t-answer I am interested in here. Standard set theory takes as a primitive the understanding presupposed in the non-t answer. I suggest that we could also take the concept of containment presupposed in the t-answer as primitive.

Note that "t-member" is not mereological, at least if mereology is more than just assumptions about the transitivity, reflexivity and symmetry properties of the "part of" relation. If we take mereology to entail that a whole is defined solely in terms of its ultimate, atomic parts then the t-membership based system is not mereology.

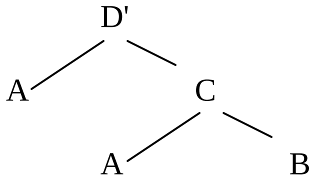
6. Move is Undefineable

Suppose we have a category C that immediately dominates categories A and B. Consider then the categories D, D' and D'' where D immediately dominates C and nothing else, D' immediately dominates A and C and nothing else while D'' immediately dominates B and C and nothing else. In standard tree notation we have (5)-(7), where (6) and (7) will be (8) and (9) respectively in terms of multiple domination or re-merge.

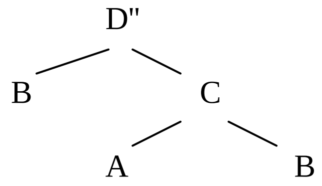
(5)



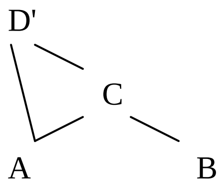
(6)



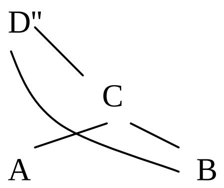
(7)



(8)



(9)



However given the notion of t-member characterized by the conditions in (3) and (4), the five intuitive characterizations of three apparent configurations by the trees above will all be formally equivalent, --the same object.

C in (5) = C in (6)/(8)= C in (7)/(9), since by the identity axiom (4), for all S, C= S iff t-member(A,S) t-member(B, S) and there is no J, such that t-member(J, S) and J is neither A or B. Next by the identity axiom again, for all S', D= S' iff t-member(A, S'), t-member(B, S'), t-member (C, S') and there is no J, such that t-member(J, S') and J is neither A or B or C. Hence D=D'=D", all trees in (5) through (9) reduce to (5).

Thus the intuitive characterization of 'move A by re-merging A (or merging A's copy) with C'corresponds to creating D together with the relations "t-member(A, D)" "t-member (C, D)" But "t-member (A, D)" adds nothing, since this follows from "t-member (C, D)" and "t-member (A, C)" by the transitivity clause in (2). In other words move or chain is undefinable in the system that uses t-membership instead of the standard it-membership of set theory. (This argument assumes of course the understanding of merge as simple (t-)set formation and the standard understanding of move or chain as remerge of a subpart S of an object O to O, where remerged S is type or token identical to S.)

While the t-set approach cannot express move or chain as standardly understood, it is clearly expressive enough in other respects. It was argued earlier (Brody 2000a, 2002) that syntactic relations are based on domination and the spec-head relation instead of c-command. But even assuming that all syntactic relations are c-command based, we can ensure that they can be captured by defining this notion. On the simplest definition x c-commands y iff x, y are t-sets, both t-members of some t-set z, and x is not a t-member of any set that y is not a t-member of. We may or may not add the usual proviso that y is not a t-member of x. (Note that this standard definition is automatically sensitive to the structural status of the deepest occurences of x and y.)

7. T-sets

It is important to see that the identity axiom (4) above is weaker than the identity axiom of standard set theory, in that there are objects like those illustrated by (5), (6) and (7) above that are identical given this condition, but not under the standard set identity condition. This does not mean that we cannot define the weaker condition on the basis of the stronger, but that is not the approach I am pursuing, where the notion of "t-member" is a primitive. The stronger standard notion of identity is not definable however solely in terms the weaker notion used here.)

The notion of "t-member" replaces the notion of set-membership. So just like an atomic element (lexical item) or a set can be a member of a set, we can define "t-set" as the object whose t-members are either atomic (lexical) or t-sets. (Putting aside here the question of empty sets and empty t-sets.) So the set S={A, {BC}} has two members A and {BC}, but due to the transitivity of the t-member relation, the t-set T=t{A, t{BC}} has four t-members A, t{BC}, B and C. Furthermore, in the more interesting case, the set S={A, {AB}} still has two members, namely A and {A,B} but the t-set T=t{A, t{AB}} has three t-members: t{AB}, A and B. The same three as T'=t{B, t{AB}} and T"=t{t{AB}}.

The proposal made here can be considered as the generalization of a standard set-theoretical observation. Take {A, A}, that is the set that results from A merged with itself. It looks different from {A}, but if the object created is anything like a set, then {A, A}={A}, --we have a single object represented (misleadingly?) in two ways. So calling {A, A} and {A}, say tree-pictures, these two tree-pictures correspond to a single syntactic object (Cf. Bury 2003 for a discussion of the relevance of this set-theoretical fact for syntax). Now let us return to the t-membership / t-set suggestion, namely that trees are not sets but t-sets. This suggestion can be thought of as extending the standard set-theoretical observation about duplicates, -- from sisters to all duplicates in a c-command relation.

If we assume "t-sets" defined in terms of "t-member", then the identity condition (3) defines the same identity classes as a definition of identity for standard sets stated in terms of a simplified (label-disregarding) version of Chomsky's (1994) notion of "term" would give us if we used it in an irreflexive fashion. In this version of Chomsky's definition, for any structure K, K is a term of K and if L is a term of K , then the members of the members of L are terms of K. Consider instead the related concept term*, such that M is a term* of K iff. M is a term of K and M is not K So if we used "term*" and defined set-identity by substituting "term*" for "t-member" then we would have a system where S={A, {AB}}={B, {AB}}={{AB}}, -- all three have three terms*: {AB}, A and B. And clearly this system is weaker than one which makes all the same distinctions and in addition distinguishes these sets also. Perhaps I should emphasize that the identity condition for sets in terms of "term*" would not be additional to the identity axiom in terms of members, but a substitute for it under this scheme.

But proceeding in this way we must introduce more complexity, we apparently need to define a notion ("term*") in addition to the simple standard notion of set. So that is why I stress that I take "t-member" to be a primitive. No complexity appears to be added this way. In other words the transitive understanding of membership is not more complex than the non-transitive. If anything, exactly the opposite: we are talking about elements of/in something and it is the non-transitive understanding that presupposes an additional locality condition.

8. Asymmetry, Transitive Sets

As pointed out by Klaus Abels (p.c.), the standard short definition of ordered pairs $<x,y>=\{x, \{x,y\}\}$ would fail if construed in terms of t-membership, as $<x,y>=t\{x, t\{x,y\}\}$. For example: $<x,y>=t\{x, t\{x,y\}\}= t\{y, t\{x,y\}\}= <y,x>$. Indeed, although the argument does not carry over directly to the longer Kuratovsky pair $<x,y>=\{\{x\}, \{x,y\}\}$ — $t\{t\{x\}, t\{x,y\}\} \neq t\{t\{y\}, t\{x,y\}\}$— this definition would also fail the ordered pair property that $y\neq z \supset <x,y>\neq<x,z>$. With $y\neq z$ and $x=t\{y,z\}$, $<x,y> = t\{t\{t\{y,z\}\}, t\{t\{y,z\},y\}\} = t\{t\{t\{y,z\}\}, t\{t\{y,z\},z\}\} = <x,z>$. We apparently need to restrict our ordered pairs $<x,y>$ to cases where x is not a t-member of y and y is not a t-member of x. A more general definition that dispenses with this restriction is proposed by Abels (2005).

According to von Neumann the number $n+1 = (n \cup \{n\})$; according to Zermelo $n+1 = \{n\}$. Sidestepping the more general philosophical problems of what numbers are (Benacerraf 1965, Steinhart 2002), note that it is standard set theory that creates the distinction between these two natural identifications of numbers with sets. In the weaker system of the t-set approach we cannot distinguish the two characterizations, --as above $\{n \cup \{n\}\}$ is equivalent to $\{\{n\}\}$, -- just like under standard set theory $\{n,n\}=\{n\}$. In general, since in the t-set approach $x \in y \rightarrow x \subset y$ (by transitivity (2) above), all t-sets will be transitive sets. (A is a transitive set iff $x \in A \rightarrow x \subseteq A$.)

Notes

* I am very grateful to Klaus Abels, Noam Chomsky, George Galfalvi, László Kálmán and Hans van der Koot for helpful comments and discussions.

[1]To the best of my recollection essentially this argument was made repeatedly by Chomsky during a workshop he held at the Budapest Tilt in 2004. A less surprising version of the claim sometimes attributed to Chomsky is that given merge, there is no a priory reason to restrict it to non-overlapping objects. Hence given merge, move must be possible in the default case, unless specifically exluded (though not necessarily a distinct operation). As I shall argue below not even this weaker implication is a purely logical necessity. (It is true however, given set-theory.)

Michael Brody
December 2004,
(slightly revised: July 2005)

References

Abels, K., 2005. untitled. University of Tromso
Benacerraf, Paul. (1965) What numbers could not be. *Philosophical Review*. **7**4. 47-73.

Boeckx, Cedric and Norbert Hornstein (2003) “Reply to Control is not movement” *Linguistic Inquiry* 34. 269-280.
Brody, Michael (1995). *Lexico-Logical Form: A Radically Minimalist Theory*. Cambridge, Mass.:MIT Press.
Brody, Michael (1997). Perfect chains. In L. Haegeman (ed.). *Elements of Grammar*. 139-167. Amsterdam: Kluwer.
Brody, Michael (1998). The minimalist program and a perfect syntax. *Mind and Language* 13. 205-214.
Brody, Michael (2000a) Mirror Theory: Syntactic Representation in Perfect Syntax. *Linguistic Inquiry* 31.1. 29-56
Brody, Michael (2000b) Relating Syntactic Elements. *Syntax* 3.
Brody, Michael (2001) One More Time. *Syntax* 4.2. pp.126-138
Brody, Michael (2002) On the status of derivations and representations. in Samuel Epstein and Daniel Seely eds. *Derivation and Explanation in the Minimalist Program*. Blackwell Publishers.
Bury, Dirk. *Phrase Structure and Derived Heads*. Doctoral Dissertation. UCL. London
Chomsky, Noam (1994). Bare Phrase Structure. *MIT Occasional Papers in Linguistics*.
Hornstein, Norbert (1998). Movement and chains. *Syntax* 1. 99-127.
Hornstein, Norbert (2000). On A-chains. A reply to Brody. *Syntax* 3.
Landau, Idan. 2003. Movement out of Control. *Linguistic Inquiry* 34.3. 471-498.
Nunes, Jairo (2001), Sideward movement. *Linguistic Inquiry* 32.2. 303-344.
Nunes, Jairo (2004), *Linearizaton of Chains and Sideward movement.* Cambridge, Mass.:MIT Press.
Steinhart, Eric (2002) Why Numbers Are Sets. *Synthese* 133. 3. 343 – 361.