On Minimal Search

Alan Hezao Ke, University of Michigan, Ann Arbor

Introduction: Chomsky (2013, 2015) suggests that labeling and agree can be reduced to minimal search (MS), a 3rd factor domain general search algorithm. However, a serious implementation of this appealing idea requires a formal definition of MS, which has not been provided. I propose a formalization of MS that defines it as a search algorithm in labeling and agree. This formalization shows that labeling and agree can only be partially reduced to MS. The empirical gains are illustrated with subject-complementizer agreement in Lubukusu and other agreement phenomena.

Proposal: (1) formalizes a definition of MS, with an example of agreement shown in (2), where uφ and vφ stand for unvalued and valued φ-features, respectively.

(1) MS=<SA, SD, ST>, where SA=search algorithm, $SD_{\in sets}$=search domain (the domain that SA operates on), ST = search target (the features that SA looks for).

Search Algorithm (SA):

a. Given SD and ST, inspect every head member of SD, looking for a match for the ST.

b. If ST is found, return the heads (feature bundles) bearing ST and go to Step (c); Otherwise, get the set members in SD and store them as list L

  i. If L is empty, search fails and go to Step (c); otherwise

  ii. assign each of the sets in L as SD and go to Step (a) for all SDs in parallel

c. Terminate search

(2) α = {$be_{[u\varphi]}$, {$_\beta$ {$a_{[v\varphi]}$, {*little*, *boy*}}, {*kicking,* {*the*, *ball*}}}} (given (1a))

1. SD=β, ST=uφ → ST not found in β 2. L=[{$a_{[v\varphi]}$, {*little*, *boy*}}, {*kicking,* {*the*, *ball*}}]

3. SD_1={$a_{[v\varphi]}$, {*little*, *boy*}}, SD_2={*kicking,* {*the*, *ball*}}

4. SD_1, ST → find vφ on $a_{[v\varphi]}$; SD_2, ST → ST not found 5. Return $a_{[v\varphi]}$; terminate search

In (2), MS looks into its SD β to find heads bearing the ST uφ (Step 1). Two set members are found (Step 2) and they are assigned as new SDs (Step 3) of new MSs (Step 4). Finally, the head bearing the vφ are found and returned (Step 4 and 5).

Theoretical Implications: (i) This definition unifies the SAs in agree and labeling; the only difference between them concerns the SD and ST. For the labeling of α in (3a), ST=any head, SD is the set to be labeled, i.e. α, and X and Y are returned by MS; for the agreement between X and Y in β (3b), ST=unvalued features (uFs) on X, and SD is the sister of the head bearing ST (to be revised blow). Y, which bears matching valued features (vFs), will be returned.

(3) a. α = {{$_{XP}$ X, JP}, {$_{YP}$ Y, KP}} b. β ={$X_{[uFs]}$, {$_\delta$ …, {$_{YP}$ $Y_{[vFs]}$, ZP}}}

(ii) (1) also shows that labeling and agree cannot be completely reduced to MS. Although MS presupposes an SD and ST, the specific values of SD and ST are assigned by labeling and agree. Labeling and agree are thus linguistic mechanisms that are independently needed. (iii) (1) implies that the label of a finite TP with a subject should not be the ordered pair of features in agreement, i.e. <φ,φ>, as was previously assumed (e.g. Chomsky 2013). If *a little boy* in (2) moves to TP Spec, and we apply MS for labeling, with SD=TP and ST=any head, *a* and *be* will be returned, rather than <φ,φ>. If we would want to return <φ,φ>, we must instead set ST=φ. However, such a labeling algorithm is not the one we use to label other sets such as the VP {*kick*, {*the*, *ball*}}, where we have ST=any head. A unified MS for labeling will therefore return only heads.

Empirical Consequences: The above definition of MS can partially unify agree and labeling because the values of SD and ST are independently assigned by relevant operations. This unique feature has far-reaching empirical consequences. Due to space limitations, I present below only its implications for the subject-complementizer (S-C) agreement in Lubukusu.

S-C agreement: Diercks (2013) argues convincingly that in Lubukusu an embedded C agrees only with the subject, not the object, of a next higher clause in their φ-features. Diercks claims that such subject-orientation in agree exemplified in (4) (analyzed in 5, cf. Carstens 2016) can be explained if we assume a null reflexive in SpecCP which agrees with C *li*, and the reflexive raises to adjoin to a higher T where it is bound by the matrix subject (cf. Chomsky 1986).

(4) Ba-ba-ndu ba-bol-el-a Alfredi ba-li a-kha-khil-e.

2-2-people 2S-said-AP-FV 1Alfred 2-that 1S-FUT-conquer
'The people told Alfred that he will win.'

(5) [$_{CP}$ C ba-ba-ndu$_j$ T [$_{vP}$ t$_{j[v\varphi]}$ *v* [$_{VP}$ V [$_{DP}$ Alfredi] [$_{ForceP}$ REFL$_j$ ba-li$_{[u\varphi]}$ [$_{IntP}$ Int [$_{FinP}$ Fin [$_{TP}$ ...]]]]]]]

This analysis correctly captures the lack of intervention effects with regard to the object *Alfredi*. A regular upward agree analysis (e.g. Zeijlstra 2012), by contrast, incorrectly predicts that C should instead agree with the object. However, a potentially serious drawback of Diercks' analysis is that a null reflexive is positioned in every embedded CP, yet they make no semantic (or phonological) contribution to their local CP. They are not logophors or elements that are related to any thematic roles, and they receive no identifiable thematic interpretation themselves.

Proposed Analysis: This study instead argues that "subject-orientation" in (4) can be derived by MS defined in (1) with an independently assigned SD. I adopt that agree proceeds in two cycles, an assumption that is independently needed to address cross-linguistic cyclic agree (e.g. Béjar and Rezac 2009, Keine and Dash 2017). I also assume that MS observes Chomsky's (2001) Phase Impenetrability Condition, that is, a phase head complement is transferred and is not accessible to MS when a higher phase head enters the derivation. Cyclic agree implemented with MS can be demonstrated with (5). When the C$_{Force}$ *li*, bearing uφ, merges with the Int(errogative)P, MS (Cycle One) is conducted, taking this IntP as its SD and the uφ as its ST. This is a typical case of MS discussed in (2). This MS fails as the lower subject has been transferred after two phase heads, Fin(ite) and Int, merge in, following Carstens (2016). Only when first cycle MS fails to value the uFs will a second cycle MS (below) be executed. The derivation continues until the lexical array of the next phase is exhausted, that is, when the *v*P-internal subject merges in. Another MS is conducted (Cycle Two) with ST=uφ and SD=the syntactic object that is built so far, i.e. the matrix *v*P. This MS finds first the *v*P-internal subject *babandu*, which causes intervention to the object *Alfredi*. It is thus the Cycle Two MS that accounts for the subject-orientation phenomenon in (4).

Advantages of the Current Analysis over Diercks' (2013): (i) The MS-based analysis is simpler and more intuitive as the subject-C agreement is established directly by agree, whereas Diercks needs to assume two distinct agreement relations, subject-reflexive and null reflexive-C dependency. The subject-reflexive binding dependency has been argued NOT to be a primitive operation but one that is reducible to agree (e.g. Reuland 2011). In addition, the null reflexive-C dependency analysis requires the reflexive to be involved in φ-feature agreement, an agreement that has been shown to be impossible (so called "anaphor (non-)agreement effects", cf. Preminger 2018). (ii) The current analysis can be extended to other cases of upward agree whereas Diercks' analysis cannot. For instance, multiple agree in Japanese and negative concord in Czech (6a, b), with the analyses in (7a, b), can be considered a case of MS in Cycle Two. In (7a, b), the uNEG and uCase features cannot be valued by Cycle One MS as the valuers are at higher positions. Cycle Two MSs will be conducted when the lexical arrays of the matrix CPs are exhausted, with ST=uNEG/ uCase and SD=CP, thus correctly establishing agreement relations by returning the operator bearing vNEG (7a) and the T head bearing vCase (7b). The reanalysis of these important cases of upward agree as Cycle Two MS, together with the preferable analysis it provides for the S-C agreement in Lubukusu (and many other phenomena that we cannot discuss here due to space limitations), suggest that the MS based agree system is an empirically viable alternative to upward agree. As a result, agree can be unified as an MS-based operation as defined in (1).

(6) a. Dnes nikdo *(ne)volá nikomu (Zeijlstra 2012:501)
Today n-body NEG.calls n-body
'Today nobody is calling anybody.'

b. John-ga [yosouijouni nihonjin-ga eigo-ga hidoku] kanji-ta. (Hiraiwa 2001:76)
John.NOM than.expected the.Japanese.NOM English.NOM bad.INF think-PAST
'It seemed to John that the Japanese are worse at speaking English than he had expected.'

(7) a. [$_{CP}$ C Op$_{\neg[vNEG]}$ [$_{TP}$ nikdo$_{[uNEG]}$ nevolá$_{[uNEG]}$ nikomu$_{[uNEG]}$]

b. [$_{CP}$ C [$_{TP}$ T$_{[vCase]}$ [$_{vP}$ DP$_{[uCase]}$ … [$_{TP}$ DP$_{[uCase]}$ … [… DP$_{[uCase]}$ …]]]]]