

Graph theoretic phrase markers ('trees') vs. set theoretic phrase markers

Consider the following 2 toy  $\Sigma, F$  grammars and the sentence (terminal string)  $df$

Grammar 1	Grammar 2
$\Sigma: A$	$\Sigma: A$
F:	F:
$A \rightarrow BE$	$A \rightarrow CE$
$B \rightarrow C$	$C \rightarrow B$
$C \rightarrow d$	$B \rightarrow d$
$E \rightarrow f$	$E \rightarrow f$

All the equivalent derivations of  $df$  by Grammar 1:

A	A	A
B E	B E	B E
C E	C E	B f
C f	d E	C f
d f	d f	d f

Graph theoretic phrase marker ('reduced derivation tree') of  $df$  via Grammar 1:



Set theoretic phrase marker of  $df$  via Grammar 1:

$\{A, BE, Bf, CE, Cf, dE, df\}$

All the equivalent derivations of  $df$  by Grammar 2:

A	A	A
C E	C E	C E
B E	B E	C f
B f	d E	B f
d f	d f	d f

Graph theoretic phrase marker ('reduced derivation tree') of  $df$  via Grammar 2:



Set theoretic phrase marker of  $df$  via Grammar 2:

$\{A, BE, Bf, CE, Cf, dE, df\}$

Note that the trees are different, but the sets are the same. The trees contain more information. (Can you see what the extra info is?) If Chomsky is correct that all a phrase marker needs to do is determine the *is a* relations between portions of the terminal string and single non-terminal symbols, then the set theoretic representation suffices. And if it suffices, it is preferable. Why provide extra information that is never used?