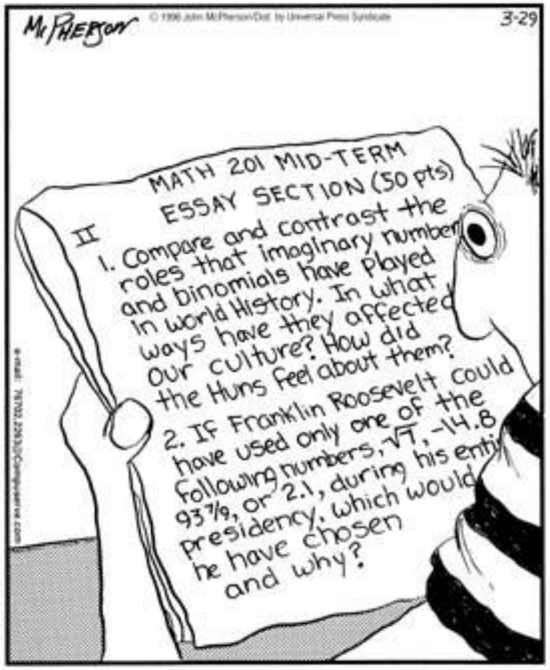
Let L=

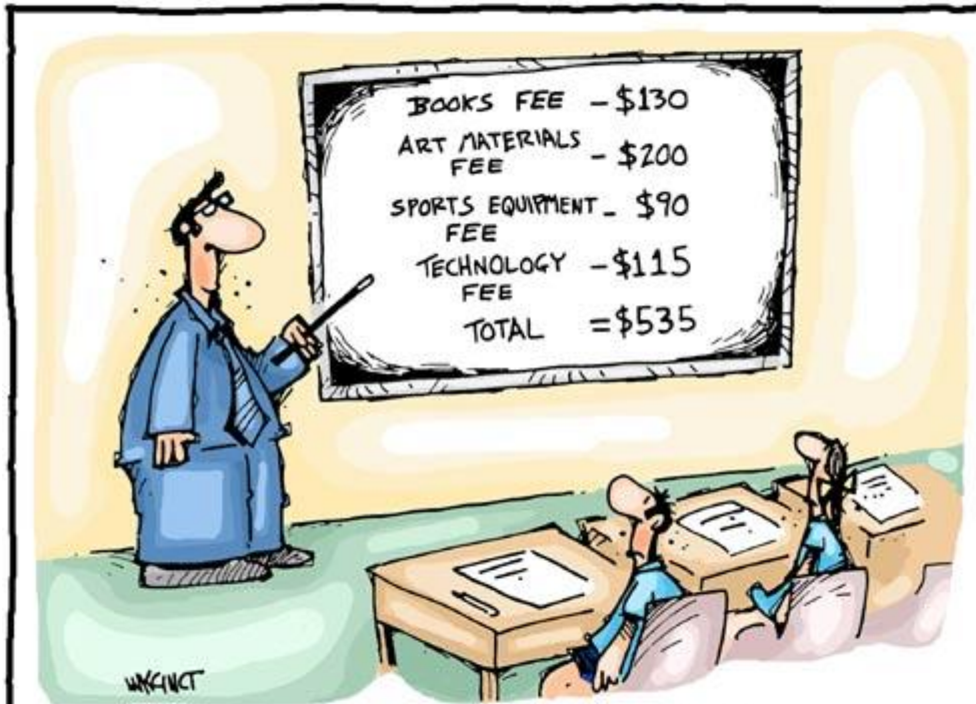
$\{ w = u\, v : u \in \{a, b\}^*, v \in \{c, d\}^* \text{ and } |u| = |v| \}$

1. Design a context-free grammar that generates L.
2. Use your grammar and the construction from last class to design a PDA that accepts L.

CLOSE TO HOME JOHN McPHERSON



A diabolical new testing technique: math essay questions.

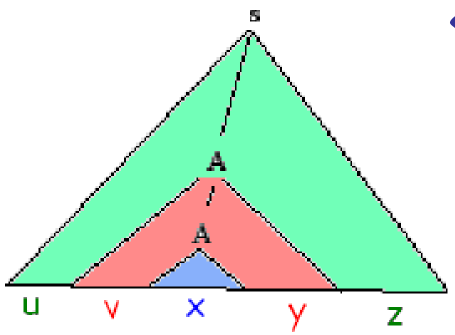


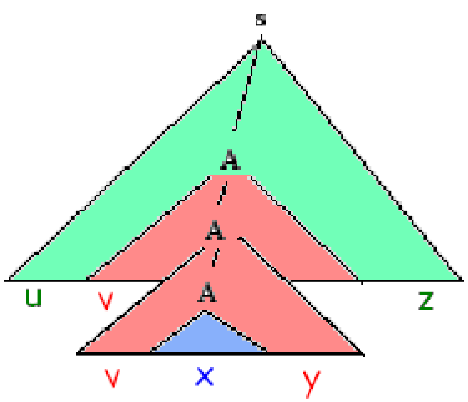
Teaching mathematics in government schools today.

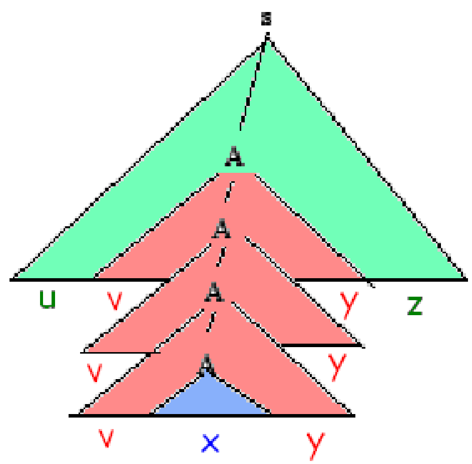
29/01 2007-050 © John Ditchburn

Languages which are not context-free

$\{ a^n b^n c^n : n \geq 0\}$







"But on the box it says, 'Open here'."

Languages which are not context-free

The pumping theorem, a statement about context-free languages, is used to create proofs (by contradiction) that languages are not context-free.

Closure properties can be used: We will prove next class that intersecting a context-free language and a regular language gives a context-free language.

Once we have some languages proven to not be context-free, we can give counterexamples to show that context-free languages are not closed under intersection or complement.

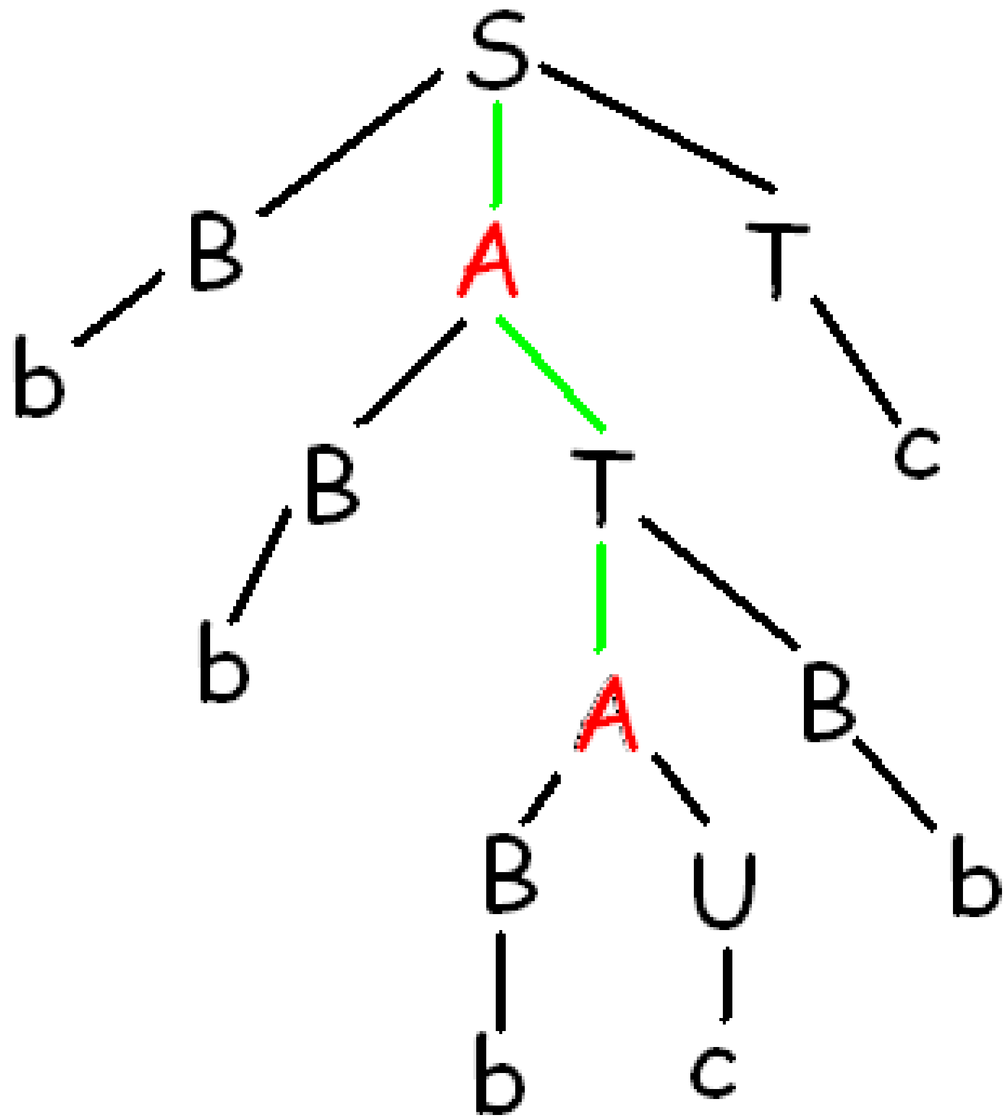
The Pumping Theorem for Context-Free Languages:

Let G be a context-free grammar.

Then there exists some constant k which depends on G such that for any string w which is generated by G with $|w| \geq k$,

there exists u, v, x, y, z, such that

1. $w = u\ v\ x\ y\ z$,
2. $|v| + |y| \geq 1$, and
3. $u\ v^n\ x\ y^n\ z$ is in L for all $n \geq 0$.



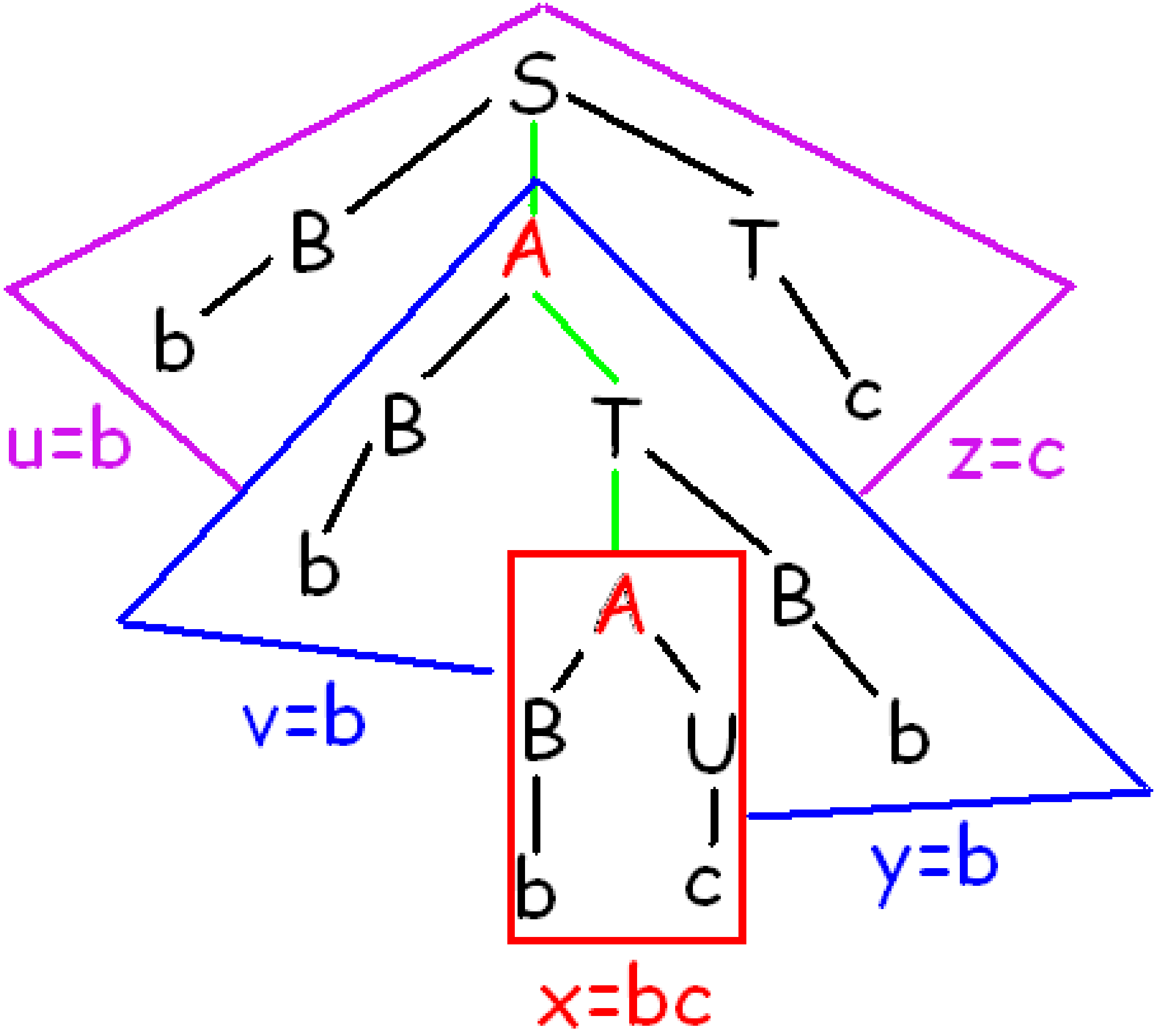
To pump:

Look for a path from root with a repeated non-terminal.

Path:

S - A - T - A

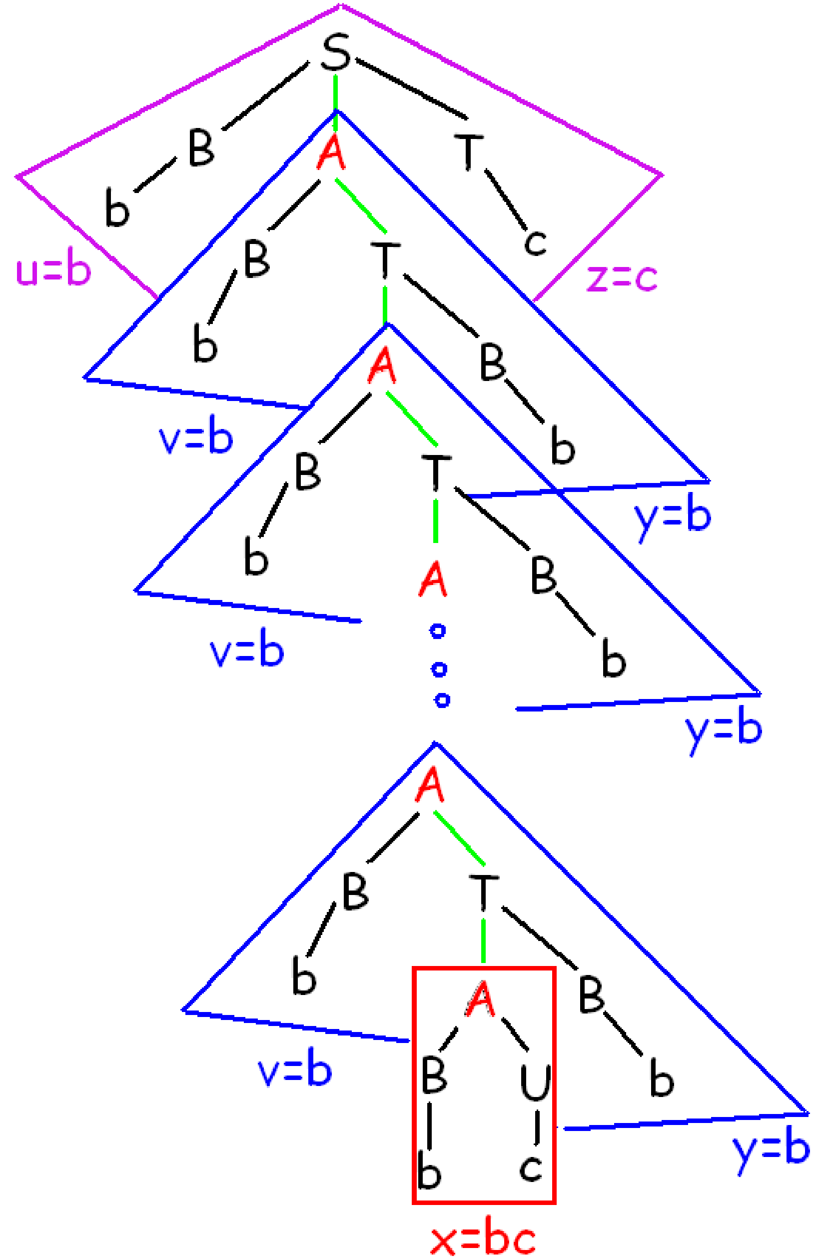
Non-terminal A is repeated.


S
B
b
A
B
b
T
A
B
U
b
c
B
b
T
c
u=b
v=b
x=bc
y=b
z=c

To pump n times:

New yield is:

$= u\, v^n\, x\, y^n\, z$



Theorem: $L = \{ a^n b^n c^n : n \geq 0\}$ is not context-free. Choose $w = a^k b^k c^k$.

Consider all possibilities for u, v, x, y, z:

Case 1: v is in the a's

Case 2: v is in the b's

Case 3: v is in the c's

Case 4: v has both a's and b's

Case 5: v has both b's and c's

Case 6: v has a's, b's and c's

$w = a^k b^k c^k$.

Case 1: v is in the a's

Case 1.1: y is in the a's

Case 1.2: y is in the b's

Case 1.3: y is in the c's

Case 1.4: y has both a's and b's

Case 1.5: y has both b's and c's

Case 1.6: y has a's, b's and c's

$w = a^k b^k c^k$.

Case 2: v is in the b's

Case 2.1: y is in the b's

Case 2.2: y is in the c's

Case 2.3: y has both b's and c's

Case 3: v is in the c's

Case 3.1: y is in the c's

$w = a^k b^k c^k$.

Case 4: v has both a's and b's

Case 4.1: y is in the b's

Case 4.2: y is in the c's

Case 4.3: y has both b's and c's

Case 5: v has both b's and c's

Case 5.1: y is in the c's

Case 6: v has a's, b's and c's

Case 6.1: y is in the c's

$w = a^k b^k c^k$. MUST CONSIDER ALL CASES:

Case 1: v is in the a's

- Case 1.1: y is in the a's
- Case 1.2: y is in the b's
- Case 1.3: y is in the c's
- Case 1.4: y has both a's and b's
- Case 1.5: y has both b's and c's
- Case 1.6: y has a's, b's and c's

Case 2: v is in the b's

- Case 2.1: y is in the b's
- Case 2.2: y is in the c's
- Case 2.3: y has both b's and c's

Case 3: v is in the c's

- Case 3.1: y is in the c's

Case 4: v has both a's and b's

- Case 4.1: y is in the b's
- Case 4.2: y is in the c's
- Case 4.3: y has b's and c's

Case 5: v has both b's and c's

- Case 5.1: y is in the c's

Case 6: v has a's, b's and c's

- Case 6.1: y is in the c's

Case 1: v is in the a's

Case 1.1: y is in the a's

Case 1.2: y is in the b's

Case 1.3: y is in the c's

Case 1.4: y has both a's and b's

Case 1.5: y has both b's and c's

Case 1.6: y has a's, b's and c's

Case 2: v is in the b's

Case 2.1: y is in the b's

Case 2.2: y is in the c's

Case 2.3: y has both b's and c's

Case 3: v is in the c's

Case 3.1: y is in the c's

CASE A: v and y contain at most one type of symbol.

CASE B: v or y has more than one type of symbol.

Case 4: v has both a's and b's

Case 4.1: y is in the b's

Case 4.2: y is in the c's

Case 4.3: y has b's and c's

Case 5: v has both b's and c's

Case 5.1: y is in the c's

Case 6: v has a's, b's and c's

Case 6.1: y is in the c's

Theorem: $L = \{ a^n b^n c^n \ : n \geq 0\}$ is not context-free. Choose $w= a^k b^k c^k$.

CASE A: v and y contain at most one type of symbol.

Pump zero times. The number of occurrences of one or two types of symbols decreases but the number of occurrences of at least one type of symbol remains the same. Hence the resulting string no longer has equal numbers of a's, b's and c's.

CASE B: v or y has more than one type of symbol.

Pump twice. The resulting string is not in L since it is not of the form a*b*c*.

Final formal proof (by contradiction):

Theorem: $L = \{ a^n b^n c^n : n \geq 0\}$ is not context-free.

Assume that L is context-free. Then there exists some constant k such that for all strings w in L with length at least k, the pumping theorem holds.

Choose $w = a^k b^k c^k$. This string w is in L and the length of w is at least k and therefore, the pumping theorem holds.

The Pumping Theorem for Context-Free Languages:

Let G be a context-free grammar.

Then there exists some constant k which depends on G such that for any string w which is generated by G with $|w| \geq k$,

there exists u, v, x, y, z, such that

1. $w = u\, v\, x\, y\, z$,
2. $|v| + |y| \geq 1$, and
3. $u\, v^n\, x\, y^n\, z$ is in L for all $n \geq 0$.

Consider all ways to factor w as uvxyz such that $|v| + |y| \geq 1$:

CASE A: v and y contain at most one type of symbol.

Pump zero times. The number of occurrences of one or two types of symbols decreases but the number of occurrences of at least one type of symbol remains the same. Hence the resulting string no longer has equal numbers of a's, b's and c's.

CASE B: v or y has more than one type of symbol.

Pump twice. The resulting string is not in L since it is not of the form a*b*c*.

Note that I am very careful with the wording of the proof. For example:

Case 1.2: v is in the a's, y is in the b's

Factorizations are of the form:

$a^i \; a^j \; a^{k-i-j} \; b^r \; b^s \; b^{k-r-s} \; c^k$ where $j + s \geq 1$.

$ \; v \qquad\qquad\qquad y$

Pumping zero times gives:

$a^i (a^j)^0 a^{k-i-j} \; b^r (b^s)^0 b^{k-r-s} \; c^k$

$= a^i a^{k-i-j} \; b^r b^{k-r-s} \; c^k$

$= a^{k-j} \; b^{k-s} \; c^k$ NOTE: $j + s \geq 1$.

If $j=0$: $a^k \; b^{k-s} \; c^k \Rightarrow$ less b's.

If $s=0$: $a^{k-j} \; b^k \; c^k \Rightarrow$ less a's.

If $j \neq 0$ and $s \neq 0$: $a^{k-j} \; b^{k-s} \; c^k \Rightarrow$ more c's than a's or b's.

Theorem: L= {w in {a, b, c}* : w has the same number of a's, b's and c's} is not context-free.

Proof (by contradiction):

Assume L is context-free. Since a context-free language intersected with a regular language must be context-free, this means that:

$L' = L \cap a^* b^* c^*$ is context-free.

But $L' = \{ a^n b^n c^n : n \geq 0\}$ and hence L' is not context-free.

The Pumping Theorem for Context-Free Languages:

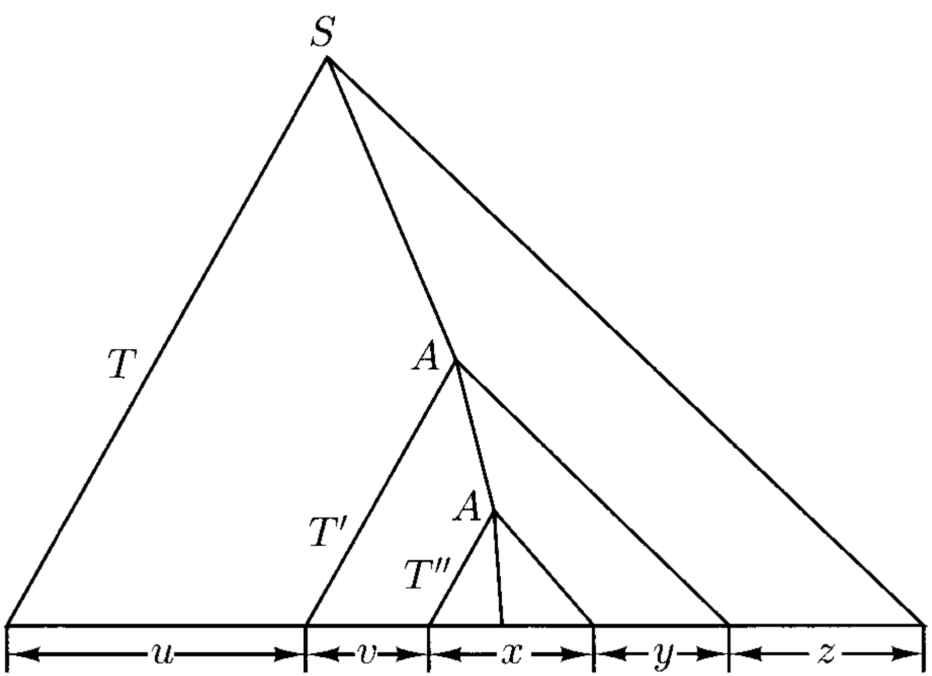
Let G be a context-free grammar.

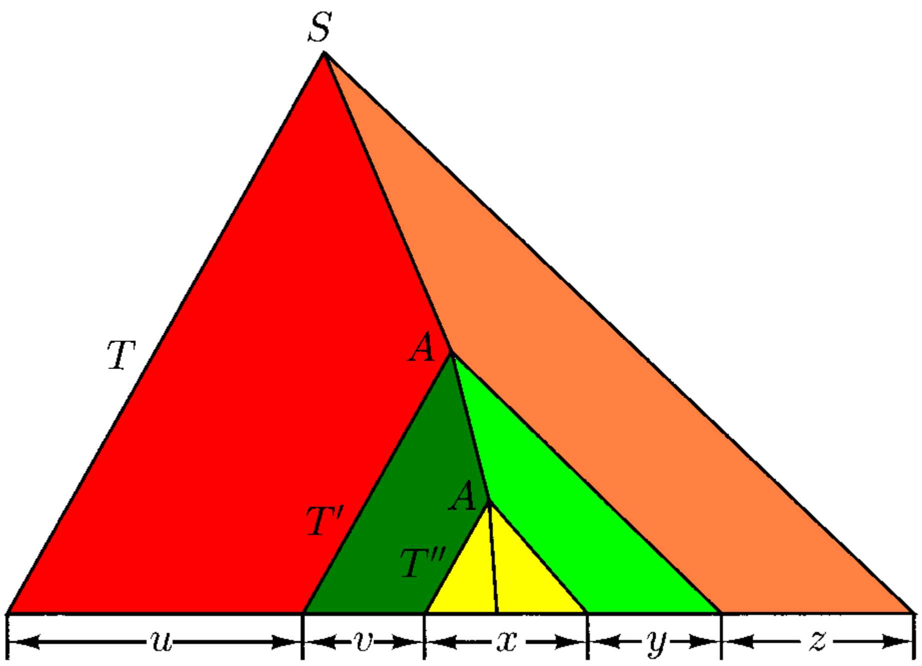
Then there exists some constant k which depends on G such that for any string w which is generated by G with $|w| \geq k$,

there exists u, v, x, y, z, such that

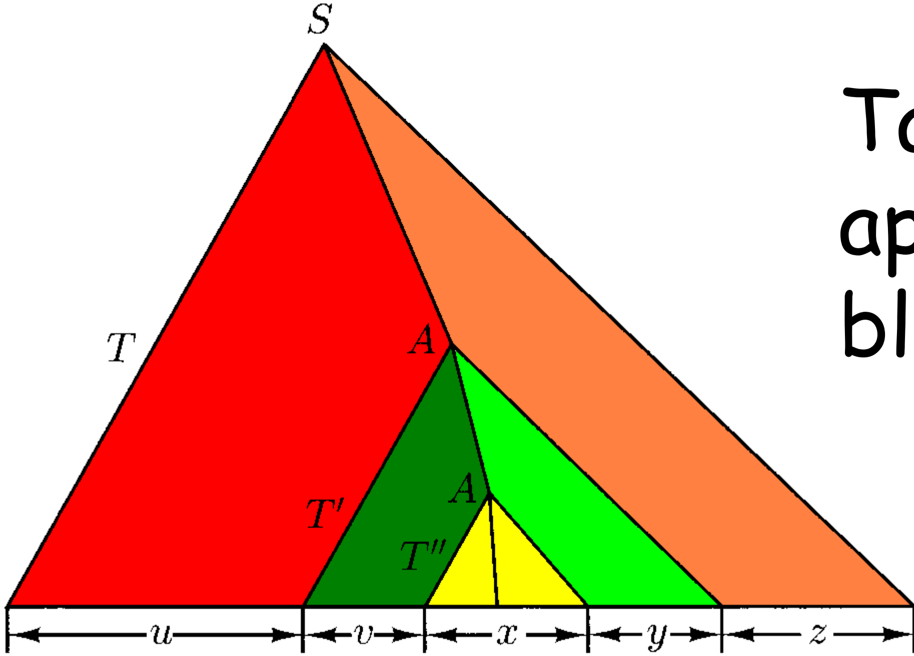
1. $w = u\,v\,x\,y\,z$,
2. $|v| + |y| \geq 1$, and
3. $u\,v^n\,x\,y^n\,z$ is in L for all $n \geq 0$.

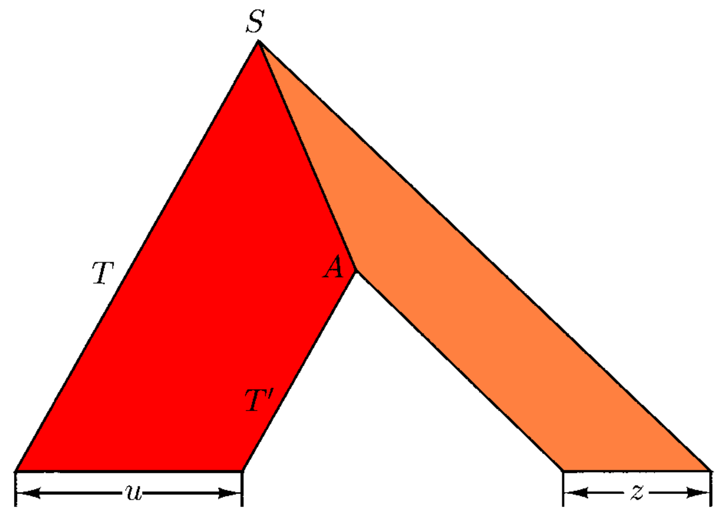
Proof: If the string w is long enough, then there is a path from the root with a non-terminal repeated which does not have both v and y equal to the empty string.

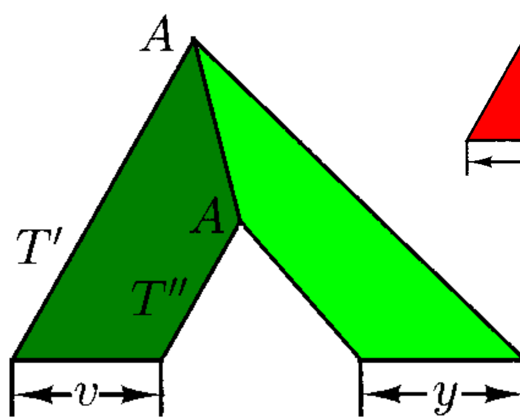


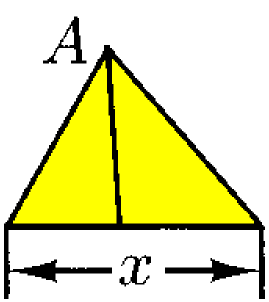


Take parse tree apart to get building blocks.

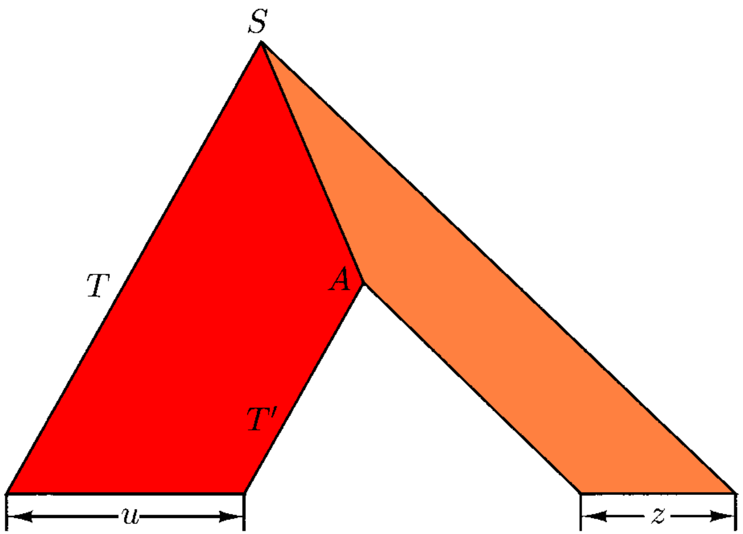


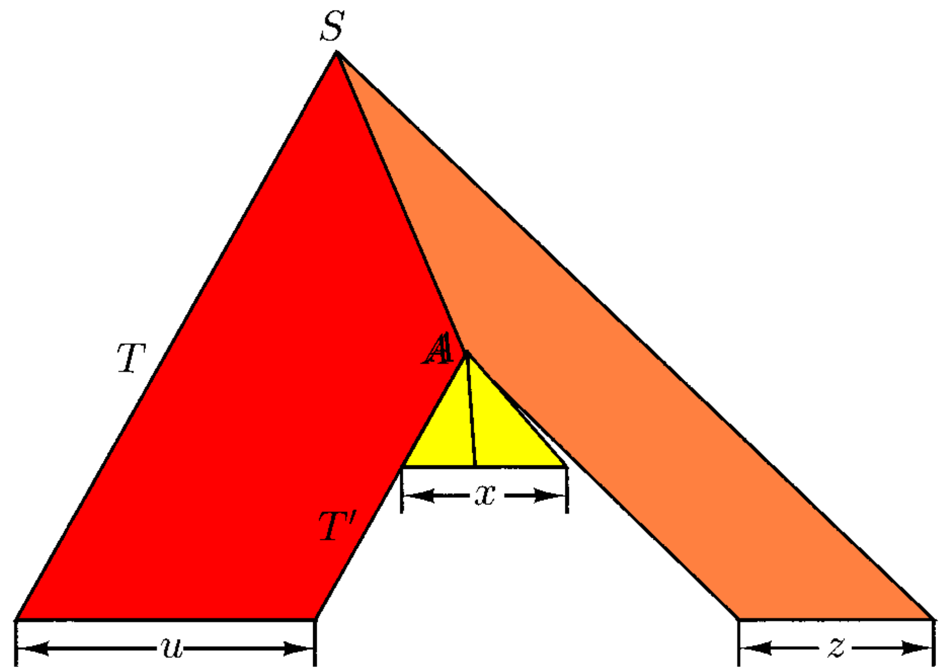


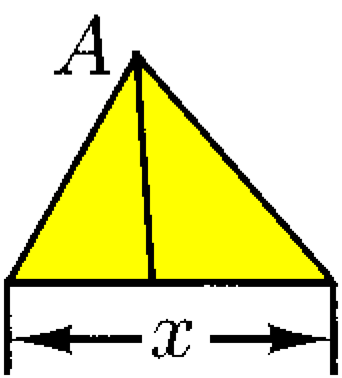




To pump 0 times:



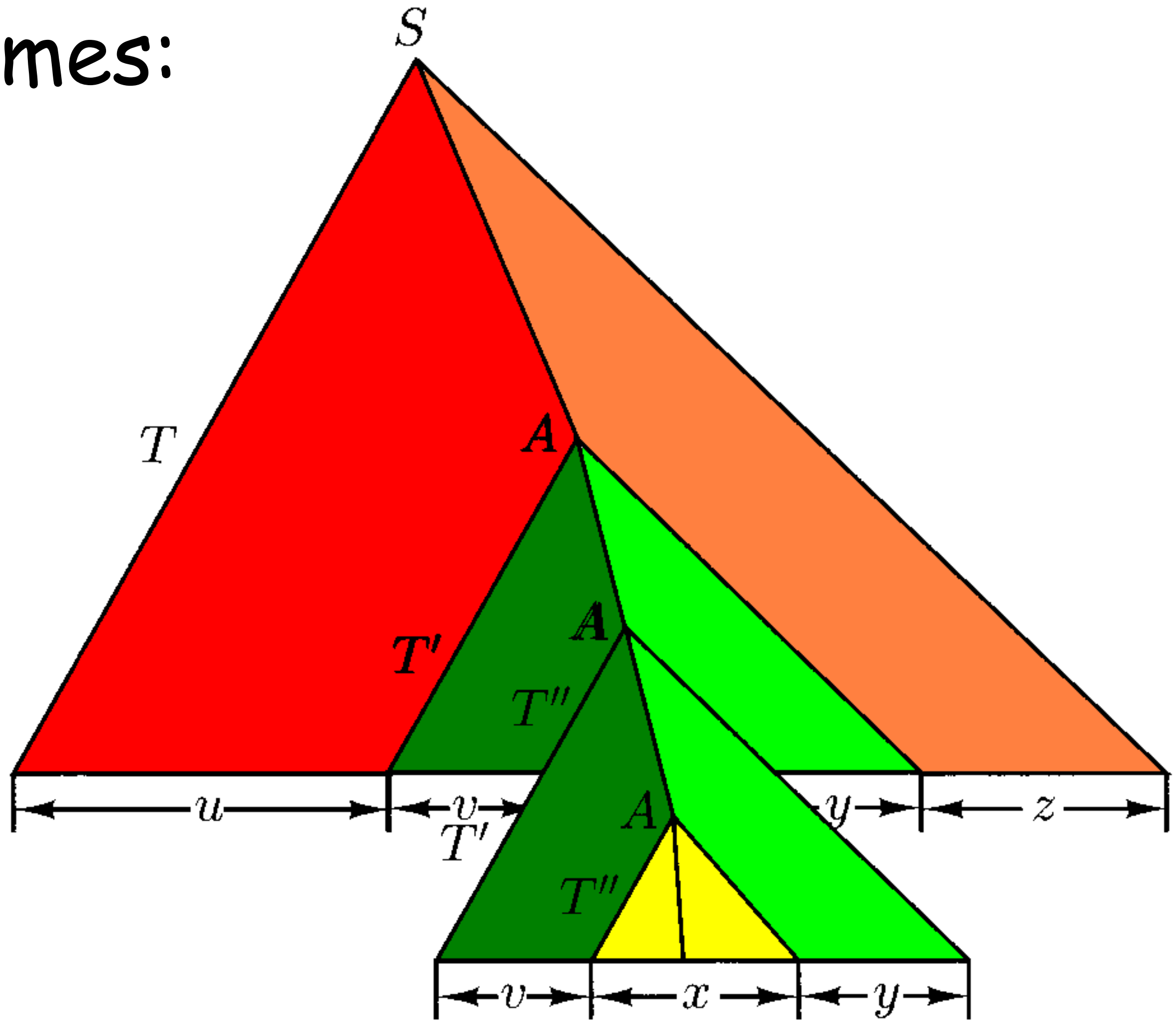




New yield is u x z

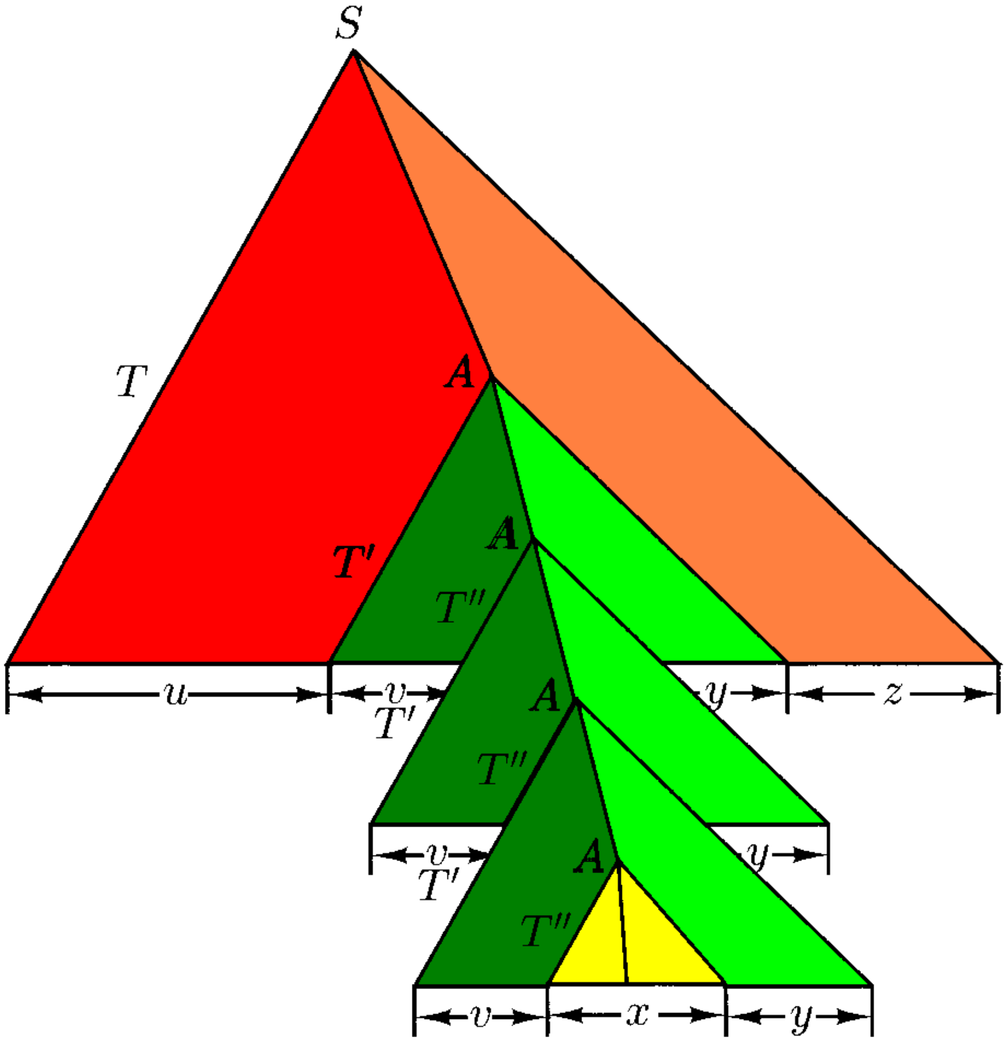
$= u\, v^0\, x\, y^0\, z$

To pump 2 times:



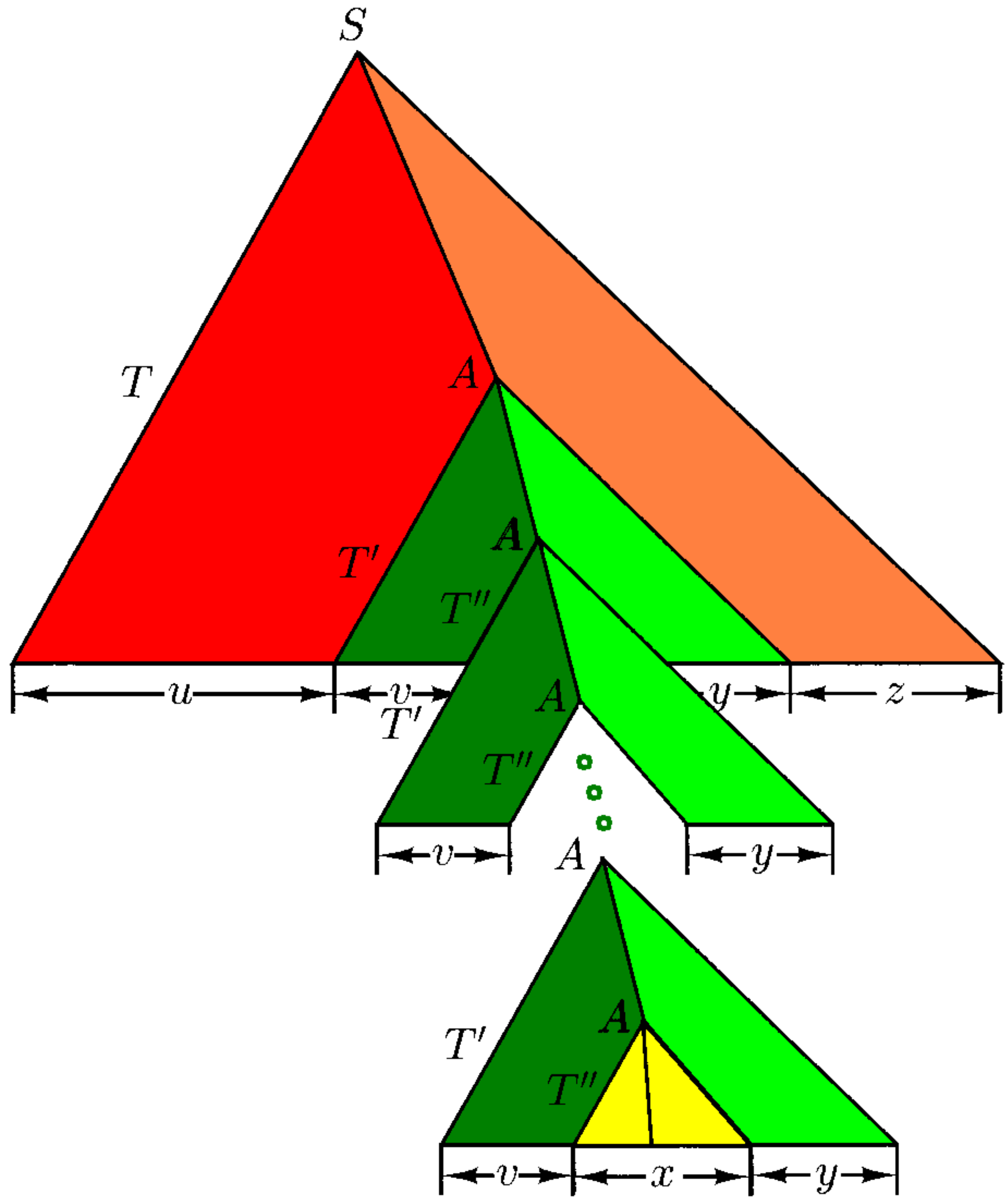
New yield is u v v x y y z = u v^2 x y^2 z

To pump 3 times:



New yield is u v v v x y y y z = u v^3 x y^3 z

To pump n times:



New yield is

$u\ v^n\ x\ y^n\ z$