Refined upper bounds on the size of the condensed neighbourhood of sequences
Cedric Chauve, May 2021.
Preliminary version: France Paquet-Nadeau, July 2017.

```
> restart:
  with(gfun):
  march('open',
  "C:\\Users\\cchau\\Desktop\\SFU\\PSC_2021\\CondensedNeighbourhoods
  \\algolib.mla");
```

libname := "C:\Program Files\Maple 2021\lib", **(1)**

"C:\Users\cchau\Desktop\SFU\PSC_2021\CondensedNeighbourhoods\algolib.mla"

Variables:
k = number of symbols in the query of the edit script
s = size of alphabet (default: 2)
d = distance bound, maximum number of edit operations in edit scripts

We use the two functions below to evaluate S(s, k, d).
The function S will be kept symbolic while GenRec will evaluate each term; S1 is used for S in GenRec.
For actual evaluation we assume that s has been given an integer value.

```
> GenRec:= proc(s, k, d::integer)
  local j:
  if d = 0 or (type(k, posint) and k <= d) then
    return 1
  else
    return S1(s, k-1, d)
    + (s-1) * S1(s, k-1, d-1)
    + (s-1) * add(s^j * S1(s, k-2, d-1-j), j=0..d-1)
    + (s-1)^2 * add(s^j * S1(s, k-2, d-2-j), j=0..d-2)
    + add(S1(s, k-2-j, d-1-j), j=0..d-1):
  fi:
  end proc:
> S:= proc(s, k, d::integer)
  local j:
  option remember;
  if d = 0 then
    return 1
  elif type(k, nonnegint) then
    if k <= d then
      return 1
    else
      return procname (s, k-1, d)
      +(s-1) * procname(s, k-1, d-1)
      +(s-1) * add(s^j * procname(s, k-2, d-1-j), j=0..d-1)
      +(s-1)^2 * add(s^j * procname(s, k-2, d-2-j), j=0..d-2)
      + add(procname(s, k-2-j, d-1-j), j=0..d-1):
    fi:
  else
    'S'(args):
  fi:
  end proc:
```

Transforming the recurrences of S into generating function.

We will denote by F[d] the gen. functions S_{s,d)(z) obtained from this method for a given value of d.
To obtain F[1] we compute S(s, k, 1) and input the resulting expression manually into ***rectodiffeq*** to solve the differential expression defined by the recurrence.
For the following values of d, we construct 'expression' with the recurrences of 'S'.

```
> F := 'F':  # Initializing
> r := 9 : # Rank  of the gen. function we comute
> eval(subs(S1 = S, GenRec(s, k, 1)))
```

$$S(s, k-1, 1) + 2s - 1 \tag{2}$$

```
> F[1]  := factor(solve(rectodiffeq({a(n)=a(n-1) + (2*s) - 1, a(0)=
  1}, a(n), f(z)), f(z)));
```

$$F_1 := \frac{2zs - 2z + 1}{(z-1)^2} \tag{3}$$

We now iterate building F[d], i.e. S_{s,d)(z), from lower values of d.
We look at the gen. functions for d from 2 to 5.

```
> d_max := 5 :
> for d from 2 to d_max do: # dPrime used as an index to build the
  table of gen. functions
    # RHS of the equation of S(s, k,d) with similar terms
    # grouped together
    expression := S(s, k, d)=collect(eval(subs(S1=S, GenRec(s, k,
  d))), S):
    # Collecting the constant terms
    const := eval(subs(S=(()-> 0), expression)):
    # Removing the constant terms from the equation to keep
    # only terms of the form 'S(s, k, d)'
    expression := expression - const;
    # Transforming 'expression' into a gen. function;
    for dd from 1 to d do
      # dd corresponds to the distance in the terms 'S(s, k, d)'
      for j from 0 to d do
        # j is the index for terms substracted to 'k'
        # Substitute the term S(s, k-j, dd) with the equation
        # for F[dd] adjusted to the index
        expression := subs(S(s, k-j, dd)=z^j * (F[dd]-(add (eval(S
   (s, l+1, dd)) * z^l, l=0..(d-j-1)))), expression):
      od:
    od:
    # Adding back the constant terms transformed
    # into generating function
    expression := expression +  const * (z^d / (1-z)):
    F[d] := factor(solve(expression, F[d]));
    print(simplify(F[d]));
  od:
```

$$\frac{-1 + z^3 + \left(-4s^2 + 4s - 3\right)z^2 + 2z}{(z-1)^3}$$

$$\frac{1}{(z-1)^4}\left(1 + \left(2s^3 - s^2 - 3s + 3\right)z^5 + \left(-4s^3 + 6s^2 - 2s - 3\right)z^4 + \left(10s^3 - 17s^2 + 11s - 2\right)z^3 + 3z^2 - 3z\right)$$

$$\frac{1}{(z-1)^5}\left(-1+\left(4 s^4-8 s^3+8 s^2-6 s+3\right) z^7+\left(-20 s^4+32 s^3-14 s^2+4 s-6\right) z^6+\left(28 s^4-56 s^3+40 s^2-14 s+8\right) z^5+\left(-28 s^4+64 s^3-58 s^2+24 s-7\right) z^4+4 z^3-6 z^2+4 z\right)$$

$$\frac{1}{(z-1)^6}\left(1+\left(14 s^5-29 s^4+20 s^3-5 s^2+1\right) z^9+\left(-72 s^5+164 s^4-140 s^3+52 s^2-9\right) z^8+\left(156 s^5-354 s^4+290 s^3-93 s^2-9 s+20\right) z^7+\left(-152 s^5+380 s^4-368 s^3+170 s^2-30 s-10\right) z^6+\left(86 s^5-241 s^4+278 s^3-164 s^2+49 s-4\right) z^5+5 z^4-10 z^3+10 z^2-5 z\right) \tag{4}$$

We now define T_{s,d}(z) parameterized by d as we will consider s as fixed and z is the formal variable of the gen. functions.
We show the value of T_{s,d}(z) for d from 1 to 5.

```
> for d from 1 to d_max do:
    T[d] := F[d] + add(s^j * z * (F[d-j]-1), j=1..d-1) + (s^d / (1 - z)):
    print(simplify(T[d]))
  od:
```

$$\frac{(s-2) z+s+1}{(z-1)^2}$$

$$\frac{-s z^4+\left(2 s^2+s+1\right) z^3+\left(-7 s^2+4 s-3\right) z^2+\left(2 s^2+2\right) z-s^2-1}{(z-1)^3}$$

$$\frac{1}{(z-1)^4}\left(\left(2 s^3-2 s^2-3 s+3\right) z^5+\left(-6 s^3+12 s^2-2 s-3\right) z^4+\left(9 s^3-22 s^2+10 s-2\right) z^3+\left(5 s^3+s+3\right) z^2+\left(-3 s^3-3\right) z+s^3+1\right)$$

$$\frac{1}{(z-1)^5}\left(\left(6 s^4-9 s^3+5 s^2-3 s+3\right) z^7+\left(-26 s^4+38 s^3-13 s^2-3 s-6\right) z^6+\left(40 s^4-72 s^3+53 s^2-8 s+8\right) z^5+\left(-37 s^4+70 s^3-70 s^2+19 s-7\right) z^4+\left(6 s^4+5 s^3+2 s^2+4 s+4\right) z^3+\left(-8 s^4-s^2-s-6\right) z^2+\left(4 s^4+4\right) z-s^4-1\right)$$

$$\frac{1}{(z-1)^6}\left(\left(18 s^5-37 s^4+28 s^3-11 s^2+3 s+1\right) z^9+\left(-94 s^5+203 s^4-165 s^3+65 s^2-9 s-9\right) z^8+\left(196 s^5-435 s^4+348 s^3-121 s^2+4 s+20\right) z^7+\left(-190 s^5+478 s^4-454 s^3+221 s^2-39 s-10\right) z^6+\left(93 s^5-283 s^4+311 s^3-199 s^2+45 s-4\right) z^5+\left(15 s^5-s^4+14 s^3+9 s^2+10 s+5\right) z^4+\left(-14 s^5-5 s^4-3 s^3-5 s^2-5 s-10\right) z^3+\left(12 s^5+s^3+s^2+s+10\right) z^2+\left(-5 s^5-5\right) z+s^5+1\right) \tag{5}$$

Evaluation for increasing values of s

```
> s := 2:
```

First we compute the generating functions for S(s, k, d) for each value of d.

```
> for d from 1 to d_max do
    Fn[d, s] := equivalent(F[d], z, k);
    print(factor(numer(F[d])), factor(denom(F[d])));
  od:
```

Warning, (in anonymous procedure created in infsing) environment variable `_EnvAllSolutions` declared as a global variable
Warning, (in anonymous procedure created in infsing) try statement has no catch or finally clause

$$2z+1,\ (z-1)^2$$

$$z^3-11z^2+2z-1,\ (z-1)^3$$

$$9z^5-15z^4+32z^3+3z^2-3z+1,\ (z-1)^4$$

$$23z^7-118z^6+140z^5-127z^4+4z^3-6z^2+4z-1,\ (z-1)^5$$

$$125z^9-601z^8+1278z^7-1118z^6+558z^5+5z^4-10z^3+10z^2-5z+1,\ (z-1)^6 \quad \textbf{(1.1)}$$

Now we compute the asymptotic of S(s, k, d).

```
> seq(print(map(z->ifactor(numer(z))/ifactor(denom(z)), coeff
  (convert(Fn[d, s] * d!, polynom), k, d)) * k^d), d=1..d_max);
```

$$(3)\,k$$

$$(3)^2\,k^2$$

$$(3)^3\,k^3$$

$$(3)^4\,k^4$$

$$(3)^5\,k^5 \quad \textbf{(1.2)}$$

For the same value of s, the generating function T_{s,d}(z) and the asymptotics of CN(s,k,d).

```
> for d from 1 to d_max do
    Tn[d,s]:=equivalent(T[d], z, k);
    print(factor(numer(T[d])), factor(denom(T[d])));
  od:
```

$$3,\ (z-1)^2$$

$$-2z^4+11z^3-23z^2+10z-5,\ (z-1)^3$$

$$5z^5-7z^4+2z^3+45z^2-27z+9,\ (z-1)^4$$

$$41z^7-176z^6+268z^5-281z^4+156z^3-140z^2+68z-17,\ (z-1)^5$$

$$171z^9-847z^8+1640z^7-1268z^6+226z^5+637z^4-592z^3+408z^2-165z+33,\ (z-1)^6 \quad \textbf{(1.3)}$$

```
> seq(print(map(z->ifactor(numer(z))/ifactor(denom(z)), coeff
  (convert(Tn[d,s]*d!, polynom), k, d))*k^d), d=1..d_max);
```

$$(3)\,k$$

$$(3)^2\,k^2$$

$$(3)^3\,k^3$$

$$(3)^4\,k^4$$

$$(3)^5\,k^5 \quad \textbf{(1.4)}$$

```
> s:=3:
> for d from 1 to d_max do
    Fn[d,s]:=equivalent(F[d], z, k);
    print(factor(numer(F[d])), factor(denom(F[d])));
  od:
```

$$4z+1,\ (z-1)^2$$
$$z^3-27z^2+2z-1,\ (z-1)^3$$
$$39z^5-63z^4+148z^3+3z^2-3z+1,\ (z-1)^4$$
$$165z^7-876z^6+1082z^5-997z^4+4z^3-6z^2+4z-1,\ (z-1)^5$$
$$1549z^9-7533z^8+16220z^7-14662z^6+7550z^5+5z^4-10z^3+10z^2-5z+1,\ (z-1)^6 \quad \textbf{(2.1)}$$

```
> seq(print(map(z->ifactor(numer(z))/ifactor(denom(z)), coeff
  (convert(Fn[d,s]*d!, polynom), k, d))*k^d), d=1..d_max);
```

$$(5)\,k$$
$$(5)^2\,k^2$$
$$(5)^3\,k^3$$
$$(5)^4\,k^4$$
$$(5)^5\,k^5 \quad \textbf{(2.2)}$$

```
> for d from 1 to d do
    Tn[d,s]:=equivalent(T[d], z, k);
    print(factor(numer(T[d])), factor(denom(T[d])));
  od:
```

$$z+4,\ (z-1)^2$$
$$-3z^4+22z^3-54z^2+20z-10,\ (z-1)^3$$
$$30z^5-63z^4+73z^3+141z^2-84z+28,\ (z-1)^4$$
$$282z^7-1212z^6+1757z^5-1687z^4+655z^3-666z^2+328z-82,\ (z-1)^5$$
$$2044z^9-10305z^8+20732z^7-17848z^6+6413z^5+4058z^4-3958z^3+2965z^2-1220z$$
$$+244,\ (z-1)^6$$
$$T_6,\ 1 \quad \textbf{(2.3)}$$

```
> seq(print(map(z->ifactor(numer(z))/ifactor(denom(z)), coeff
  (convert(Tn[d,s]*d!, polynom), k, d))*k^d), d=1..d_max);
```

$$(5)\,k$$
$$(5)^2\,k^2$$
$$(5)^3\,k^3$$
$$(5)^4\,k^4$$
$$(5)^5\,k^5 \quad \textbf{(2.4)}$$

```
> s:=4:
```

```
> for d from 1 to d_max do
    Fn[d,s]:=equivalent(F[d], z, k);
    print(factor(numer(F[d])), factor(denom(F[d])));
  od:
```

$$6\,z+1,\ (z-1)^2$$

$$z^3-51\,z^2+2\,z-1,\ (z-1)^3$$

$$103\,z^5-171\,z^4+410\,z^3+3\,z^2-3\,z+1,\ (z-1)^4$$

$$619\,z^7-3286\,z^6+4176\,z^5-3911\,z^4+4\,z^3-6\,z^2+4\,z-1,\ (z-1)^5$$

$$8113\,z^9-39881\,z^8+86176\,z^7-79330\,z^6+41728\,z^5+5\,z^4-10\,z^3+10\,z^2-5\,z+1,\ (z-1)^6 \tag{3.1}$$

```
> seq(print(map(z->ifactor(numer(z))/ifactor(denom(z)), coeff
  (convert(Fn[d,s]*d!, polynom), k, d))*k^d), d=1..d_max);
```

$$(7)\,k$$

$$(7)^2\,k^2$$

$$(7)^3\,k^3$$

$$(7)^4\,k^4$$

$$(7)^5\,k^5 \tag{3.2}$$

```
> for d from 1 to d do
    Tn[d,s]:=equivalent(T[d], z, k);
    print(factor(numer(T[d])), factor(denom(T[d])));
  od:
```

$$2\,z+5,\ (z-1)^2$$

$$-4\,z^4+37\,z^3-99\,z^2+34\,z-17,\ (z-1)^3$$

$$87\,z^5-203\,z^4+262\,z^3+327\,z^2-195\,z+65,\ (z-1)^4$$

$$1031\,z^7-4450\,z^6+6456\,z^5-6043\,z^4+1908\,z^3-2074\,z^2+1028\,z-257,\ (z-1)^5$$

$$10589\,z^9-53853\,z^8+109716\,z^7-97878\,z^6+39680\,z^5+16189\,z^4-15918\,z^3+12382\,z^2-5125\,z+1025,\ (z-1)^6$$

$$T_6,\ 1 \tag{3.3}$$

```
> seq(print(map(z->ifactor(numer(z))/ifactor(denom(z)), coeff
  (convert(Tn[d,s]*d!, polynom), k, d))*k^d), d=1..d_max);
```

$$(7)\,k$$

$$(7)^2\,k^2$$

$$(7)^3\,k^3$$

$$(7)^4\,k^4$$

$$(7)^5\,k^5 \tag{3.4}$$