§1 KODA-RUSKEY INTRO 1

(Downloaded from https://cs.stanford.edu/~knuth/programs.html and typeset on May 28, 2023)

1. Intro. This program contains two implementations of the Koda-Ruskey algorithm for generating all ideals of a given forest poset [Journal of Algorithms 15 (1993), 324-340]. The common goal of both implementations is, in essence, to generate all binary strings  $b_0 \dots b_{n-1}$  in which certain bits are required to be less than or equal to specified bits that lie to their right. (For some values of j there is a value of k > j such that we don't allow  $b_k$  to be 0 when  $b_j = 1$ .) Moreover, each binary string should differ from its predecessor in exactly one bit position; the algorithm therefore defines a generalized reflected Gray code.

The given forest is represented by n pairs of nested parentheses. For example, ()()()() represents five independent bits, while ((((())))) represents five bits with  $b_0 \leq b_1 \leq b_2 \leq b_3 \leq b_4$ . A more interesting example, ((())(()())), represents six bits subject to the conditions  $b_0 \leq b_1$ ,  $b_1 \leq b_5$ ,  $b_2 \leq b_4$ ,  $b_3 \leq b_4$ ,  $b_4 \leq b_5$ . Each pair of parentheses corresponds to a bit that must not exceed the bit of its enclosing pair, if any, and the pairs are ordered by the appearances of their right parentheses.

The first implementation uses n coroutines, which call each other in a hierarchical fashion. The second uses multilinked data structures in a loopless way, so that each generation step performs a bounded number of operations to obtain the next element. I couldn't resist writing this program, because both implementations turn out to be quite interesting and instructive.

Indeed, I think it's a worthwhile challenge for people who study the science of computer programming to verify that these two implementations both define the same sequence of bitstrings. Even more challenging would be to derive the second implementation "automatically" from the first.

2 Intro Koda-ruskey §2

**2.** In this step we parse the forest into an array of "scopes": scope[j] is the index of the smallest descendant of node j, including node j itself.

```
\#define abort(m, i)
         { fprintf(stderr, m, argv[i]);  return -1; }
#define stacksize 100
                             /* max levels in the forest */
#define forestsize 100
                             /* max nodes in the forest */
\langle \text{Process the command line, parsing the given forest } 2 \rangle \equiv
  if (argc \neq 2 \lor argv[1][0] \neq `(`) abort("Usage: \\sum \%s_\\"nestedparens\\"\n", 0);
  for (j = k = l = 0; argv[1][k]; k++)
    if (argv[1][k] \equiv '(') {
       stack[l++] = j;
       if (l \equiv stacksize) abort("Stack_loverflow_l---_l\"%s\"_lis_ltoo_ldeep_lfor_lme!\n",1);
    } else if (argv[1][k] \equiv ')') {
       if (-l < 0) abort("Extra_right_parenthesis_in_\"%s\"!\n",1);
       scope[j++] = stack[l];
       if (j \equiv forestsize) abort("Memory_overflow_o---_o\"%s\"_ois_otoo_big!\n",1);
     else\ abort("The_forest_spec_\"%s\"_should_contain_only_parentheses!\n",1);
  if (l) abort("Missing_right_parenthesis_in_\"%s\"!\n",1);
  nn = j;
This code is used in section 1.
3. \langle \text{Global variables 3} \rangle \equiv
  int stack[stacksize];
                            /* nodes preceding each open leftparen, while parsing */
  int scope [forestsize];
                            /* table that exhibits each rightparen's influence */
  int nn;
               /* the actual number of nodes in the forest */
See also sections 11 and 17.
```

This code is used in section 1.

This code is used in section 1.

**4.** The coroutine implementation. Our first implementation uses a system of n cooperating programs, each of which represents a node in the forest. For convenience we will call the associated record a "cnode." If p points to a cnode, p-child points to the cnode representing its rightmost child, and p-sib points to the cnode representing its nearest sibling on the left, in the given forest.

Each cnode corresponds to a coroutine whose job is to generate all the ideals of the subforest it represents. Whenever the coroutine is invoked, it either changes one of the bits in its scope and returns *true*, or it changes nothing and returns *false*. Initially all the bits are 0; when it first returns *false*, it will have generated all legitimate bit patterns, ending with some nonzero pattern. Subsequently it will generate the patterns again in reverse order, ending with all 0s, after which it will return *false* a second time. Invoking it again and again will repeat the same process, going forwards and backwards, ad infinitum.

Each coroutine has the same basic structure, which can be described as follows in an ad hoc extension of C language:

```
coroutine p()
{
    while (1) {
        p¬bit = 1; return true;
        while (p¬child()) return true;
        return p¬sib();
        while (p¬child()) return true;
        p¬bit = 0; return true;
        return p¬sib();
    }
}
```

If either p-child or p-sib is  $\Lambda$ , the corresponding coroutine  $\Lambda()$  is assumed to simply return false.

Suppose  $p enthalphack{-}child()$  first returns false after it has been called r times; thus  $p enthalphack{-}child()$  generates r different patterns, including the initial pattern of all 0s. Similarly, suppose that  $p enthalphack{-}sib()$  generates l different patterns before first returning false. Then the coroutine p() itself will generate l(r+1) patterns in between the times when it returns false. The final bit pattern for p will be the final bit pattern for  $p enthalphack{-}sib$ , together with either  $p enthalphack{-}bit = 1$  and the final bit pattern of  $p enthalphack{-}child$  (if l is odd) or with  $p enthalphack{-}bit = 0$  and all 0s in  $p enthalphack{-}child$  (if l is even).

```
\langle \text{Type definitions 4} \rangle \equiv
  typedef enum {
    false, true
  } boolean:
  typedef struct cnode_struct {
                   /* either 0 or 1; always 1 when a child's bit is set */
    char bit:
    char state:
                     /* the current place in this cnode's coroutine */
                                        /* rightmost child in the given forest */
    struct cnode_struct *child;
    struct cnode_struct *sib;
                                      /* nearest left sibling in the given forest */
    struct cnode_struct *caller;
                                         /* which coroutine invoked this one */
  } cnode;
See also section 14.
```

**5.** When coroutine p calls coroutine q, it sets p-state to an appropriate number and also sets q-caller = p. Then control passes to q at the place determined by q-state.

When coroutine q wants to return a boolean value, it sets *coresult* to this value; then it passes control to p = q-caller at the place determined by p-state.

This program simulates coroutine linkage with a big switch statement. Actually the notion of "passing control" really means that we simply assign a value to the variable  $cur\_cnode$ .

The value of q-caller for every cnode q is completely determined by the structure of the given forest, so we could set it once and for all during the initialization instead of setting it dynamically as done here. But what the heck.

```
#define cocall(q, s)
          { cur\_cnode \neg state = s;
            if (q) q \rightarrow caller = cur\_cnode, cur\_cnode = q;
            else coresult = false;
            goto cogo; }
#define bitchange(b, s)
          { cur\_cnode \neg bit = b, coresult = true; coreturn(s); }
\#define coreturn(s)
          { cur\_cnode \neg state = s, cur\_cnode = cur\_cnode \neg caller;}
            goto cogo; }
\langle Repeatedly switch to the proper part of the current coroutine 5\rangle \equiv
cogo: switch (cur_cnode→state) {
     (Cases for coroutine states 6);
  default: abort("%s:_UNknown_state_code_(this_can't_happen)!\n",0);
This code is used in section 9.
6. In its initial state 0, a coroutine turns its bit on, returns true, and enters state 1.
\langle Cases for coroutine states _{6}\rangle \equiv
case 0: bitchange(1,1);
See also sections 7, 8, and 12.
This code is used in section 5.
```

7. The purpose of state 1 is to run through all bit patterns of the current node's children, starting with all 0s and ending when they reach their final pattern. At that point we invoke the current node's nearest left sibling and enter state 3. An intermediate state 2 is defined for the purpose of examining the result after calling the child coroutine.

The purpose of state 3 is simply to return to whoever called us, passing along the information in *coresult*, which tells whether any of our left siblings has changed one of its bits. Then we will continue in state 4.

```
⟨ Cases for coroutine states 6⟩ +≡
case 1: cocall(cur_cnode¬child, 2);
case 2: if (coresult) coreturn(1);
  cocall(cur_cnode¬sib, 3);
case 3: coreturn(4);
```

8. State 4 is rather like state 1, except that the child coroutine is now running through its bit patterns in reverse order. Finally it reduces them all to 0s, and returns *false* the next time we attempt to invoke it. At that point we reset the current bit, return *true*, and enter state 6.

State 6 invokes the sibling coroutine, leading to state 7. And state 7 is like state 3, but it takes us back to state 0 instead of state 4.

```
⟨ Cases for coroutine states 6⟩ +≡
case 4: cocall(cur_cnode¬child, 5);
case 5: if (coresult) coreturn(4);
bitchange(0,6);
case 6: cocall(cur_cnode¬sib, 7);
case 7: coreturn(0);
```

**9.** Hey, the implementation is done already, except that we have to get it started and write the code that controls it at the outermost level.

```
⟨ Generate the strings with a coroutine implementation 9⟩ ≡
   {
    register cnode *cur_cnode;
    ⟨ Initialize the cnode structure 10⟩;
    ⟨ Repeatedly switch to the proper part of the current coroutine 5⟩;
    }
This code is used in section 1.
```

10. We allocate a special cnode to represent the external world outside of the given forest.

```
#define root\_cnode\ cnode\_table[nn].child \langle Initialize the cnode structure 10 \rangle \equiv scope[nn] = 0; for (k = 0;\ k \leq nn;\ k++) if (scope[k] < k) \{ cnode\_table[k].child = cnode\_table + k - 1; for (j = k - 1;\ scope[j] > scope[k];\ j = scope[j] - 1)\ cnode\_table[j].sib = cnode\_table + scope[j] - 1; \} cur\_cnode = cnode\_table + nn; goto upward\_step; This code is used in section 9.
```

11. ⟨Global variables 3⟩ +≡
 cnode cnode\_table[forestsize + 1]; /\* the cnodes \*/
 boolean coresult; /\* value returned by a coroutine \*/

6

12. States 8 and greater are reserved for the external (outermost) level, which simply invokes the coroutine for the entire forest and prints out the results, until the bit patterns have been generated in both the forward and reverse directions.

```
Case for coroutine states 6⟩ +≡
case 8: if (coresult) {
   upward_step: ⟨Print out all the current cnode bits 13⟩;
     cocall(root_cnode,8);
   }
   printf("..._and_now_we_generate_them_in_reverse:\n");
   goto downward_step;
case 9: if (coresult) {
   downward_step: ⟨Print out all the current cnode bits 13⟩;
     cocall(root_cnode,9);
   }
   break;

13. ⟨Print out all the current cnode bits 13⟩ ≡
   for (k = 0; k < nn; k++) putchar('0' + cnode_table[k].bit);
   putchar('\n');
This code is used in section 12.</pre>
```

14. The loopless implementation. Our coroutine implementation solves the generation problem in a nice and natural fashion, but it can be inefficient if the given forest has numerous nodes of degree one. For example, a one-tree forest like  $((\ldots()\ldots))$  with n pairs of parentheses will need approximately  $\binom{n}{2}$  coroutine invocations to generate n+1 bitstrings.

Our second implementation reduces the work in such cases to O(n); in fact, it needs only a bounded number of operations to generate each bitstring after the first. It does, however, need a slightly more complex data structure with four link fields.

The basic idea is to work with a dynamically varying list of nodes called the current *fringe* of the forest. The fringe consists of all node whose bit is 1, together with their children. We maintain it as a doubly linked list, so that p-left and p-right are the neighbors of p on the left and right. A special node head is provided to make the list circular; thus head-right and head-left are the leftmost and rightmost fringe nodes.

A fringe node is said to be said to be either active or passive. Every node is active when it joins the fringe, but it becomes passive for at least a short time when its bit changes value; at such times the node is essentially shifting direction between going forward or backward, as in the coroutine implementation. (A passive node corresponds roughly to a coroutine that is asking its siblings to make the next move.) We save time jumping across such call-chains by using a special link field called the focus: If p is a passive fringe node whose righthand neighbor p-right is active, p-focus is the rightmost active node to the left of p in the fringe; otherwise p-focus = p. (The special focus head node is always considered to be active, for purposes of this definition, but it is not strictly speaking a member of the fringe.)

The loopless implementation works with records called lnodes, just as the coroutine implementation worked with cnodes. Besides the dynamic *bit* and *left* and *right* and *focus* fields already mentioned, each lnode also has a static field called *lchild*, representing its leftmost child. (There is no need for an *rchild* field, since p-rchild = p - 1 when p- $lchild \neq \Lambda$ .)

If p is not in the fringe, p-focus should equal p. Also, p-left and p-right are assumed to equal the nearest siblings of p to the left and right, respectively, if such siblings exist; otherwise p-left and/or p-right are undefined.

```
⟨Type definitions 4⟩ +≡
typedef struct lnode_struct {
  char bit; /* either 0 or 1; always 1 when a child's bit is set */
  struct lnode_struct *left, *right; /* neighbors in the forest and/or fringe */
  struct lnode_struct *lchild; /* leftmost child */
  struct lnode_struct *focus; /* red-tape cutter for efficiency */
} lnode;
```

**boolean** been\_there\_and\_done\_that;

Here now is the basic outline of the loopless implementation:  $\langle$  Generate the strings with a loopless implementation  $_{15}\rangle \equiv$ register lnode \*p, \*q, \*r;(Initialize the lnode structure, putting all roots into the fringe 16); **while** (1) { ⟨ Print out all the current lnode bits 22⟩;  $\langle$  Set p to the rightmost active node of the fringe, and activate everything to its right 18 $\rangle$ ; if  $(p \neq head)$  { if  $(p \rightarrow bit \equiv 0)$  { /\* moving forward \*/  $p \rightarrow bit = 1;$ (Insert the children of p after p in the fringe 19); } else {  $p \rightarrow bit = 0;$ /\* moving backward \*/  $\langle$  Delete the children of p from the fringe  $20 \rangle$ ; } else if (been\_there\_and\_done\_that) break; else {  $printf("...|and|now||we||generate||them||in||reverse:\n");$  $been\_there\_and\_done\_that = true;$  **continue**;  $\langle Make node p passive 21 \rangle;$ This code is used in section 1. 16. Initialization of the lnodes is similar to initialization of the cnodes, but more links need to be set up. #**define** head  $(lnode\_table + nn)$  $\langle$  Initialize the lnode structure, putting all roots into the fringe  $16\rangle \equiv$ for  $(k = 0; k \le nn; k++)$  {  $lnode\_table[k].focus = lnode\_table + k;$ if (scope[k] < k) { for (j = k - 1; scope[j] > scope[k]; j = scope[j] - 1) {  $lnode\_table[j].left = lnode\_table + scope[j] - 1;$  $lnode\_table[scope[j] - 1].right = lnode\_table + j;$  $lnode\_table[k].lchild = lnode\_table + j;$  $head \neg left = head - 1, (head - 1) \neg right = head;$  $head \neg right = head \neg lchild$ ,  $head \neg lchild \neg left = head$ ; This code is used in section 15. 17.  $\langle \text{Global variables } 3 \rangle + \equiv$ **lnode**  $lnode\_table[forestsize + 1];$ /\* the lnodes \*/

```
18.
      (Set p to the rightmost active node of the fringe, and activate everything to its right 18) \equiv
  q = head \rightarrow left;
  p = q \rightarrow focus;
  q \rightarrow focus = q;
This code is used in section 15.
19. (Insert the children of p after p in the fringe 19) \equiv
  if (p \rightarrow lchild) {
      q = p \rightarrow right;
      q \rightarrow left = p - 1, (p - 1) \rightarrow right = q;
     p \neg right = p \neg lchild, p \neg lchild \neg left = p;
This code is used in section 15.
20. \langle Delete the children of p from the fringe 20 \rangle \equiv
  if (p→lchild) {
     q = (p-1) \rightarrow right;
     p \neg right = q, q \neg left = p;
This code is used in section 15.
21. At this point we know that p-right is active.
\langle \text{ Make node } p \text{ passive } 21 \rangle \equiv
  p \rightarrow focus = p \rightarrow left \rightarrow focus;
  p \rightarrow left \rightarrow focus = p \rightarrow left;
This code is used in section 15.
22. \langle Print out all the current lnode bits 22 \rangle \equiv
   for (k = 0; k < nn; k++) putchar('0' + lnode\_table[k].bit);
   putchar('\n');
This code is used in section 15.
23. I used the following code when debugging.
\#define rel(f) (lnode\_table[k].f? lnode\_table[k].f - lnode\_table: -1)
\langle Print out the whole lnode structure 23\rangle \equiv
  for (k = 0; k \le nn; k++)
      printf("lnode_{\square}%d:_{\square}bit=%d,_{\square}", k, lnode_{\_}table[k].bit);
      printf("focus=%d, left=%d, left=%d, left=%d, left=%d, left=%d, rel(focus), rel(left), rel(right), rel(left));
```

10 INDEX KODA-RUSKEY  $\S24$ 

## 24. Index.

```
abort: \underline{2}, \underline{5}.
argc: \underline{1}, \underline{2}.
argv: \underline{1}, \underline{2}.
been\_there\_and\_done\_that: 15, 17.
bit: \underline{4}, 5, 13, \underline{14}, 15, 22, 23.
bitchange: 5, 6, 8.
boolean: 4, 11, 17.
caller: \underline{4}, \underline{5}.
child: \underline{4}, 7, 8, 10.
cnode: 4, 9, 11.
cnode\_struct\colon \ \underline{4}.
cnode\_table: 10, <u>11</u>, 13.
cocall: 5, 7, 8, 12.
cogo: \underline{5}.
coresult: 5, 7, 8, 11, 12.
coreturn: \underline{5}, 7, 8.
cur\_cnode: 5, 7, 8, \underline{9}, 10.
downward\_step: \underline{12}.
false: \underline{4}, 5, 8.
focus: <u>14</u>, 16, 18, 21, 23.
forestsize: \underline{2}, 3, 11, 17.
fprintf: 2.
head: 14, 15, 16, 18.
j: <u>1</u>.
k: 1.
l: <u>1</u>.
lchild: 14, 16, 19, 20, 23.
left: <u>14</u>, 16, 18, 19, 20, 21, 23.
lnode: <u>14</u>, 15, 17.
lnode_struct: \underline{14}.
lnode\_table: 16, 17, 22, 23.
main: \underline{1}.
nn: 2, \underline{3}, 10, 13, 16, 22, 23.
p: \ \ \underline{15}.
printf: 1, 12, 15, 23.
putchar: 13, 22.
q: \underline{15}.
r: \ \underline{15}.
rchild: 14.
rel: \underline{23}.
right: 14, 16, 19, 20, 21, 23.
root\_cnode: \underline{10}, \underline{12}.
scope: 2, \ \underline{3}, \ 10, \ 16.
sib: \underline{4}, 7, 8, 10.
stack: 2, \underline{3}.
stacksize: 2, 3.
state: \underline{4}, 5.
stderr: 2.
true: \underline{4}, 5, 6, 8, 15.
upward\_step: 10, \underline{12}.
```

KODA-RUSKEY NAMES OF THE SECTIONS 11

```
\langle Cases for coroutine states 6, 7, 8, 12 \rangle Used in section 5.
\langle Delete the children of p from the fringe 20\rangle Used in section 15.
\langle Generate the strings with a coroutine implementation 9\rangle Used in section 1.
Generate the strings with a loopless implementation 15 \) Used in section 1.
Global variables 3, 11, 17 Used in section 1.
(Initialize the cnode structure 10) Used in section 9.
(Initialize the lnode structure, putting all roots into the fringe 16) Used in section 15.
(Insert the children of p after p in the fringe 19) Used in section 15.
\langle Make node p passive 21 \rangle Used in section 15.
(Print out all the current cnode bits 13) Used in section 12.
(Print out all the current lnode bits 22)
                                              Used in section 15.
(Print out the whole lnode structure 23)
(Process the command line, parsing the given forest 2) Used in section 1.
(Repeatedly switch to the proper part of the current coroutine 5) Used in section 9.
(Set p to the rightmost active node of the fringe, and activate everything to its right 18) Used in section 15.
\langle Type definitions 4, 14\rangle Used in section 1.
```

## KODA-RUSKEY

	Section	Page
Intro	1	1
The coroutine implementation	4	. 3
The loopless implementation	14	. 7
Index	24	10