STRUCTURES DETECTION FOR POLYGON DECOMPOSITIONS

Martine Roussille, Véronique Bruyère, Pierre Dufour

1. Introduction.

As Keil points out in [6], the problem of decomposing a polygon into simpler components is of interest to such fields as computational geometry, syntactic pattern recognition and graphics.

We approach it in the particular framework of area calculation. We study the geometric reasoning of a human being who tries to estimate the area of a polygon with a minimal amount of calculations. Then a "right " decomposition is one that corresponds to this human criterion. Under this criterion, we realize that the use of the geometrical properties of a polygon (right angles, parallels, substructures in the shape) can produce a more suitable decomposition than algorithmic methods like [2], [4].

Our purpose is thus to develop a system that will simulate the human reasoning. For this, it will "see" the polygon and detect the geometric properties for finding the best decomposition according to our criterion.

In this paper, we study an interesting feature of the geometric reasoning namely repeated sequences in the outline of a polygon. We present our algorithms to detect them and we briefly indicate how we use their results in our decomposition system. Our algorithms generalize somewhat [1][5] to discover the partial symmetries in the shape but in opposition to [9] they also directly consider complete symmetries.

2. Notations and definitions.

For representing a polygon P, we use edges lengths and vertex external rotational angles. We describe P by a sequence of lengths and angles starting at an arbitrary vertex and travelling clockwise. It will be noted that this representation is translation-invariant and rotation-invariant. Then we encode the polygon as a circular word. For this, as in [9], we assign a unique *letter* to each different length and to each different angle. Then we translate the initial sequence of lengths and angles and produce the circular word. This word completely characterizes the polygon it is associated with. The initial problem can now be translated to *search all the non overlapping repeated factors of the circular word.* We use notations and definitions of [7].

3. Algorithms.

In a first step, we detect all the couples of non overlapping identical factors of a circular word $[w_0 \dots w_{L-1}]$ where L is the length of the word $[w]$. We introduce the predicate $REP(i,n,len)$, $(i,\ 0 < n < L,\ len \geq 1)$ which is true iff the two factors of $[w_0 \dots w_{L-1}]$ of length len which start at position i and $i+n$ are equal. For our problem, it is sufficient to detect the factors which are of maximal length. We thus introduce the following predicate to describe them :

$$\begin{aligned} MAXREP(i,n,len) \equiv\ & REP(i,n,len) \\ & \wedge\ len \leq n \\ & \wedge\ len \leq L-n \\ & \wedge\big((w_{i-1} \neq w_{i+n-1})\ \vee\ (n = len)\ \vee\ (n = L - len)\big) \\ & \wedge\big((w_{i+len} \neq w_{i+n+len})\ \vee\ (n = len)\ \vee\ (n = L - len)\big) \end{aligned}$$

where the indexes are always reduced mod L. The condition $(len \leq n)\ \wedge\ (len \leq L-n)$ expresses that the factors do not overlap

Using techniques of combinatorics on words [7], we observe properties of circular words especially for the periodicity and the overlapping factors[3]. This allows us to deduce a first algorithm for detecting all the couples of non overlapping repeated factors (cf algorithm 1).

Then we generalize to the problem of detecting all the m-uples of non overlapping identical factors of a circular word.

To generalize the precedent predicates, we define $REPM(pos,m,len)$ $m \geq 2$, $len \geq 1$ where pos is an array of size m which indicates the positions of each repeated factor

$$\begin{aligned} REPM(pos,m,len) &\equiv (\forall j,k : 1 \leq j < m, 0 \leq k < len :\ \ w_{pos[0]+k} = w_{pos[j]+k}) \\ &\equiv (\forall j : 1 \leq j < m :\ \ REP(pos[0], pos[j] - pos[0], len)) \end{aligned}$$

and $MAXREPM(pos, m, len)$ which expresses that the m non overlapping identical factors are of maximal length len but also that m is maximal (the size of the array pos may not be increased).

We have now a second algorithm from which we can deduce the useful informations concerning the periodicity and the overlapping factors.
We can also adapt the first problem to detect the couples given by a factor and its reverse *i.e.* the complete or partial palindromes.

```
var n, i, len, L : integer;
    w word;
begin
    for n = 1 to ⌊L/2⌋ do
        i ← 0;
        while i < L − 1 ∧ w_{i−1} = w_{i+n−1} do i ← i + 1;
        if i = L − 1 then < w is periodic, output the MAXREP's >;
                 else i_0 ← i;
                      repeat
                          len ← 0;
                          while w_{i+len} = w_{i+n+len} do len ← len + 1;
                          if len ≥ 1 then < output the MAXREP('s) >;
                          i ← i + len + 1;
                      until i = i_0 mod L;
        endif
    endfor
end.
```

ALGORITHM 1.

4. Decomposing with structures.

We consider the repeated factors in the word as *partial symmetries* in the shape of the associated polygon. Still we have a *complete rotational symmetry* for the polygon if the circular word is *periodic*. From these substructures of the polygon shape, we can deduce heuristic rules for finding a suitable decomposition according to the defined criterion.

Consider the $MAXREPM(pos, m, len)$ with len maximal. For instance, if we have a complete rotational symmetry then we have $m * len = L$ and we set up rules such as :

- join the $pos[i]$ positions $i = 1 \ldots m$ if $len > 2$. We then create identical components and a regular polygon.

- let $k_1, \ldots k_m$ the divisors of m, create $\frac{m}{k_i}$ identical components which will have consecutively k_i times the repeated factor. Still we will have identical components and a regular polygon.

Concerning the partial rotational symmetries, we can provide the following rules :

- join the vertices corresponding to $pos[i]$ and $pos[i]+len$ positions $\forall i \ : \ 0 \leq i \leq m$ if $len > 2$ of course. These constructions do not use the possible external angles of the factor.
- proceed in the same way with pos' such that $MAXREPM(pos', m', len')$ where $m' > m$, $2 < len' < len$. This decomposition will produce more identical components of smaller size than the first above.
- for each occurrence, try to increase len if the last object in the repetition corresponds to an angle. An interesting solution should be to draw m perpendiculars from the other terminal vertex to the edge adjacent to the terminal angle.

When we have palindromes, we say that the polygon has *axial symmetries* and we also define decomposition rules for these substructures.

5. References.

[1] S. AKL and G. TOUSSAINT, *An improved algorithm to check for polygon similarity*, Inform. Process. Lett. 7 (1978), pp. 127-128.

[2] E. L. ALLGOWER and P. H. SCHMIDT, *Computing Volumes of Polyhedra*, Math. Comp. 46 (1986), pp. 171-174.

[3] N. J. FINE and H. S. WILF, *Uniqueness theorems for periodic functions*, Proc. Am. Math. Soc. 16 (1981), pp. 109-114.

[4] M. R. GAREY, D. S. JOHNSON, F. P. PREPARATA and R. E. TARJAN, *Triangulating a simple polygon*, Informat. Proc. Lett. 7 (1978), pp. 175-179.

[5] P. T. HIGHMAN, *Optimal algorithms for finding the symmetries of a planar point set*, Inform. Process. Lett. 22 (1986), pp. 219-222.

[6] J. M. KEIL, *Decomposing a Polygon into Simpler Components*, Siam J. Comput. 14 (1985), pp. 799-817.

[7] M. LOTHAIRE, *Combinatorics on Words*, Addison-Wesley. Reading. Massachussets 1983.

[8] M. ROUSSILLE,V. BRUYERE and P. DUFOUR, *Search of repeated factors in a circular word*, to appear.

[9] S. L. TANIMOTO, *A method for detecting structure in polygons*, Pattern recognition 13 (1981), pp. 389-394.

Université de Mons-Hainaut, Faculté des Sciences
Avenue Maistriau, 15 B-7000 Mons BELGIUM
E-mail: SDUFOU @ BMSUEM11 (EARN-BITNET)