INFORMATIQUE THÉORIQUE ET APPLICATIONS

J. HIGGINS
D. CAMPBELL

Prescribed ultrametrics

Informatique théorique et applications, tome 27, nº 1 (1993), p. 1-5
<http://www.numdam.org/item?id=ITA_1993__27_1_1_0>

Informatique théorique et Applications/Theoretical Informatics and Applications
(vol. 27, n° 1, 1993, p. 1 à 5)

PRESCRIBED ULTRAMETRICS (*)

by J. HIGGINS [1] and D. CAMPBELL [1]

Communicated by G. LONGO

Abstract. – *Let $G=(S, E)$ be a subgraph of $K_n=(S, F)$, the complete graph on n vertices. Let v be a function from E to R^+. We prove two theorems on the extensibility of v. Every function v extends to a metric on F iff G is a forest. The function v extends to an ultrametric on F if and only if for all non-trivial cycles p in G,* mult *(p)>1, where* mult *(p) depends on the values of v on paths.*

Résumé. – *Soit $G=(S, E)$ un sous-graphe de $K_n=(S, F)$, le graphe complet sur n sommets. Soit v une fonction de E dans R^+. Nous prouvons deux théorèmes sur le prolongement de v. Toute fonction v se prolonge en une métrique sur F si et seulement si G est une forêt. La fonction v se prolonge en une ultramétrique sur F si et seulement si pour tout cycle non trivial p dans G, on a* mult *(p)>1, où* mult *(p) dépend des valeurs de v sur les chemins.*

INTRODUCTION

Let S be a set of points and u a non-negative real-valued function on $S\times S$. The function u is called a *metric* if

1. $u(x, y)\geqq 0$;
2. $u(x, y)=0$;
3. $u(x, y)=u(y, x)$;
4. $u(x, y)\leqq u(x, z)+u(z, y)$.

If for all z in S, u also satisfies

5. $u(x, y))\leqq \max\{u(x, y), u(z, y)\}$,

then u is called an *ultrametric*.

Ultrametrics satisfy more than the triangle inequality; inequality (5) prevents scalene triangles; that is, for any three points x, y, z of S, it is

(*) Accepted April 21, 1992.
AMS Classifications. Primary 54E35, 68R10; Secondary 05C05, 68Q25.
[1] Department of Computer Science, Brigham Young University, Provo, Utah 84602, U.S.A.

impossible that $u(x, y) < u(y, z) < u(x, z)$. To see why, note that (5) implies $u(x, z) \leq \max\{u(x, y), u(y, z)\} = u(y, z)$, a contradiction. Thus, any three points in an ultrametric space determine either an *isosceles* triangle or an *equilateral* triangle.

Ultrametrics arise in the context of p–adic evaluations on infinite fields [5]. there is interest in creating arbitrary ultrametrics on finite sets, in particular, on K_n, the complete graph on n points [1 to 4]. Since many ultrametric extensions are known to be NP-complete [3], it is most interesting that one extension can be done in a polynomial number of steps.

Theorem 1: *Let $G=(S, E)$ be a subgraph of the complete graph $K_n=(S, F)$ and let v be an arbitrary function from E to R^+. If G is a forest, then v extends to an ultrametric on F in at most $O(n^2)$ steps.*

Proof: Extend G to a spanning tree Q for K_n. Extend v to the edges of Q–G by assigning arbitrary positive number to each such edge. We use induction on n to extend v to an ultrametric u on all edges of K_n in at most $(n+1)(n-2)/2$ additional steps.

Basis: There is nothing to prove for $n=1$ or $n=2$. The case of $n=3$ is the so called isosceles restriction of an ultrametric. Namely, we define the ultrametric u on the missing edge to be the maximum of v on the other two sides. This extension takes one additional step.

Assume the result for n and consider the case $n+1$. There exists an end x of the tree Q. Let $U=S-\{x\}$. Let T be the restriction of Q to U. By induction, in at most $(n+1)(n-2)/2$ additional steps, we can find an ultrametric extension u to U of the restriction of v to T. As x is an end, there exists a unique y in U with (x, y) in Q. Let $w=v(x, y)$. For each z in $U-\{y\}$, set $u(x, z)=\max\{w, u(y, z)\}$. The number of steps to create this extension is at most $n+((n+1)(n+2)/2)=(n+2)(n+1)/2$ as claimed.

To check that our extension u is an ultrametric, we need only verify $u(a, b) \leq \max\{u(a, c), u(b, c)\}$ for all choices of distinct a, b, c in S. There are two cases: (1) x is not in $\{a, b, c\}$. (2) x is in $\{a, b, c\}$. In case (1), the inequality holds as u is an ultrametric on U. In case (2), there are two subcases: (I) y is in $\{a, b, c\}$, (II) y is not in $\{a, b, c\}$. In case (I), the inequality holds by construction. In case (II), there are three subcases: (A) $x=a$, (B) $x=b$, (C) $x=c$. Since y is not in $\{a, b, c\}$, each of these three verifications is straightforward. This concludes the proof of theorem 1.

THEOREM 2: *Let $G=(S, E)$ be a subgraph of the complete graph $K_n=(S, F)$. Then the following are equivalent:*

(*a*) *Every function $v : E \to R^+$ extends to a metric on F;*

(*b*) *G is a forest.*

Proof: Theorem 1 proves that (1 *b*) implies (1 *a*). To show (1 *a*) implies (1 *b*) it suffices to prove that if G is not a forest, then there exists a function v from E to R^+ that does *not* extend to a metric on F. If G is not a forest, then G contains a (simple) cycle $e_1, e_2, \ldots, e_k$, $k>2$. Define v on e_i, $1\leq i<k$, to be arbitrary positive numbers. Define v on the edge e_k to be any number greater than the sum of $v(e_i)$, $1\leq i<k$. Since v fails to satisfy the triangle inequality on the edge e_k, no extension of v can be a metric on F. This concludes the proof of theorem 2.

We now extend theorem 2 to ultrametrics. We will see that whether a particular function $v : S \to R^+$ has an ultrametric extension depends on the behaviour of v on non-trivial cycles of G. A cycle is any sequence of edge connected vertices $v_0 \ldots v_n$, $v_0=v_n$, allowing repeated vertices and repeated edges. A cycle is trivial, by definition, if it is a cycle with only two edges.

Let p be a (not necessarily simple) path in G. Let max (p) denote the largest value of v on p. Let mult (p) denote the number of times v attains max (p) on p. Clearly, for all paths p, mult $(p)\geq 1$.

We require two preliminary lemmas.

LEMMA 3: *A symmetric function $u : S\times S-\{(s, s) : s \text{ is in } S\} \to R^+$ is an ultrametric if and only if for each triple x, y, and z of distinct members of S,* mult $(xyzx)>1$.

Proof: If u is an ultrametric, then as remarked at the start of the paper, every triangle is either isosceles or equilateral, that is, mult $(xyzx)>1$. Conversely, to show that u must be an ultrametric when mult $(xyzx)>1$ on all triangles, it suffices to observe that (5) always holds.

LEMMA 4: *Let $G=(S, E)$ be a subgraph of the complete graph $K_n=(S, F)$. Let x and y belong to S. Let v be an arbitrary function from E to R^+. Let Q be the set of all paths from x to y in G. Let P be the set of all paths p in Q such that* mult $(p)=1$. *If all non-trivial cycles p in G satisfy mult $(p)>1$, then*

(1) *For any p_1 and p_2 in P,* max $(p_1)=$max (p_2).

(2) *For each q in Q and each p in P,* max $(q)\geq$max (p).

Proof: We prove (1) by contradiction. Suppose there were elements p_1 and p_2 of P with $\max(p_1) < \max(p_2)$. Since $c = p_1 p_2^{-1}$ is a non-trivial cycle in G, we have by hypothesis mult $(c) > 1$. Thus, there are at least two places that p_2 takes on its max, contrary to p_2 belonging to P. This proves (1). Similar proof holds for (2).

THEOREM 3: *Let $G=(S, E)$ be a subgraph of the complete graph $K_n=(S, F)$. A function $v : E \to R^+$ extends to an ultrametric on F if and only if*

($\star$) *for all non-trivial cycles p in G,* mult $(p) > 1$.

Proof: First assume that v extends to an ultrametric on F, but that ($\star$) fails for some non-trivial cycle $p=x_0 \ldots x_n$. Of all cycles p with mult $(p)=1$, choose one whose lengyh, n, is minimal. By lemma 3, mult $(p)>1$ on all 3–edged cycles. Therefore, n must be >3. Without loss of generality, let $w = \max(p) = v(x_0, x_1)$. Since mult $(p)=1$, $v(x_1, x_2)$ must be strictly less than w. Applying lemma 3 to $x_0\, x_1\, x_2\, x_0$, and knowing that $v(x_0, x_1)=w$ and $v(x_1, x_2)<w$, we conclude that $v(x_0, x_2)$ must also be w. Now form the cycle $q=x_0\, x_2 \ldots x_n$ of length $n-1$. Since mult $(q)=1$ we have obtained a contradiction to the choice of n.

Conversely, suppose that ($\star$) holds. To prove that v extends to an ultrametric, we consider two cases: G is complete, G is not complete. If G is complete, and ($\star$) holds for all triangles of G, then by lemma 3, v must be an ultrametric on S. On the other hand, if G is not complete, then there are x and y in S for which (x, y) is not in E. Let J be the union of E and the edge (x, y) and let $H=(S, J)$. Proceeding by induction on the cardinality of E, it suffices to show that H satisfies ($\star$).

Let Q be the set of paths p from x to y in G. Let P be the set of paths in Q such that mult $(p)=1$. By lemma 4,

(1) for any p_1 and p_2 in P, $\max(p_1)=\max(p_2)$;

(2) for all q in Q and all p in P, $\max(q) \geqq \max(p)$.

Define v on the edge (x, y) to be $\min\{\max(q) : q \text{ in } Q\}$. We need only show that the extension v from J to R^+ still satisfies ($\star$).

Let $s=x_0 \ldots x_n$ be a non-trivial cycle in H. Since G satisfies ($\star$) there is nothing to prove unless the edge (x, y) belongs to the cycle s. Therefore, without loss of generality, we may take $y=x_0$ and $x=x_1$. Thus, $q=x_1 \ldots x_n$, a path x to y, belongs to Q. By the definition of $v(x, y)$ and the choice of w, $v(x, y)=w \leqq \max(q)$. There are two possibilities: mult $(q)>1$, mult $(q)=1$. If mult $(q)>1$, then mult $(s)>1$ and we are done. If mult $(q)=1$, then q belongs to P. By (2) and the construction, max (q) must itself be w. Since $v\ (x_0, x_1)$ is also w, we can conclude in this case also that mult $(s)>1$. This completes the proof of theorem 3.

Theorem 2 and 3 differ significantly in computational requirements. Testing for a forest can be done in a polynomial number of steps; testing $(\star)$ for all cycles may require a factorial number of steps. For example, consider the complete graph on n vertices with a few edges removed. Such a graph has more than $n!$ non-trivial cycles.

The authors wish to thank the referee for theorem 3.

REFERENCES

1. M. Aschbacher, P. Baldi, E. Baum and R. Wilson, Embeddings of ultrametric Spaces in Finite Dimensional Structures, *S.I.A.M. J. Algebra Disc. Math.*, 1987, **8**, pp. 564-587.
2. V. Z. Feinberg, Finite Ultrametric Spaces, *Dokl. Akad. Nauk S.S.S.R.*, 1972, **202**, pp. 775-778.
3. M. Krv'anek, The Complexity of Ultrametric Partitions on Graphs, *Inform. Process. Lett.*, 1988, **27**, pp. 265-270.
4. N. Parga and M. Virasoro, The Ultrametric Organization of Memories in A Neural Network, *J. Physique*, 1986, **47**, pp. 1857-1864.
5. A. C. M. Van Rooij, Non Archimedean Functional Analysis, *Marcel Dekker*, New York, 1978.