

# Paul Erdős, 1913–1996

VLADIK KREINOVICH

# Пауль Эрдеши, 1913–1996

В. КРЕИНОВИЧ

## 1. Paul Erdős

Paul Erdős died on September 20, 1996, at the age of 83. Erdős was the most prolific of all mathematicians; he has published more mathematical papers than anyone else: more than 1200 papers with more than 300 co-authors. This phenomenon gave rise to the so-called “Erdős number”: Erdős himself had Erdős number 0, his co-authors get Erdős number 1, co-authors of his co-authors are assigned number 2, etc. It is sometimes claimed that most mathematicians have Erdős number smaller than 10.

Born in Budapest, Paul Erdős spent all his life as a wandering scholar, constantly in travel. He died in Warsaw, where he was attending a minisemester for Combinatorics; after that, he was planning to go to Vilnius for the Kubilius Conference.

Erdős's main interest was in combinatorics, including combinatorics of graphs, number theory, set theory, geometry, and mathematical analysis. His papers are widely cited and used. Among many other ideas that Erdős has pioneered is the use of probabilistic methods in combinatorics: in many cases, it is difficult to explicitly construct an object with a given property, but it is much easier to prove that the probability of a random object to have this property is positive.

Erdős's impressive productivity and tirelessness in promoting mathematics made him one of the most famous mathematicians in the world: suffice it to say that his obituary was placed on the front page of the New York Times, an honor reserved only for super-celebrities.

Erdős usually formulated his results as pure mathematical theorems, without emphasizing their possible applications; however, many of his results found important applications in various areas ranging from voting procedures in social sciences to experiment design in biology to complexity of computer algorithms.

In particular, several of his papers are related to (generalized) interval computations. Two examples:

## 2. Erdős's results related to (generalized) interval computations: first example

Measurements are never absolutely precise; expert estimates also never lead to precise values. As a result, after each measurement or estimate, we do not get an exact value of the desired quantity  $x$ ; we get the *set*  $X$  of possible values. In the case when we simply measure this quantity and we know the upper bound  $\Delta$  on the accuracy of the measuring instrument, then

this set  $X$  is an *interval*  $[\bar{x} - \Delta, \bar{x} + \Delta]$ , where  $\bar{x}$  is the measurement result. In more complicated cases, we may have more complicated sets: e.g., if we know that  $x^2$  belongs to the interval  $[1, 4]$ , then the set of possible values of  $x$  is a union of two intervals  $[-2, -1] \cup [1, 2]$ .

If we knew the exact values of several quantities  $x_1, \dots, x_n$ , then we would be able to tell which of these values are equal and which are not, i.e., for each  $i$  and  $j$ , we would know whether  $x_i = x_j$  or  $x_i \neq x_j$ . As a result, in this hypothetic case, we could divide  $n$  quantities into equivalence classes. We can describe these equivalence classes graphically if we draw  $n$  points corresponding to  $n$  quantities, and connect the points that correspond to equal quantities by edges. Then, we will get a graph that it is a union of disconnected complete subgraphs representing equivalence classes of equal quantities.

In real life, we do not know the exact values, we only know the sets  $X_1, \dots, X_n$  of possible values. In this case, for each  $i$  and  $j$ , we cannot tell whether the values of  $x_i$  and  $x_j$  are equal or not; we can only distinguish between the cases when it is *impossible* for  $x_i$  and  $x_j$  to be equal (i.e., when  $X_i \cap X_j = \emptyset$ ), and when it is *possible* for them to be equal (i.e., when  $X_i \cap X_j \neq \emptyset$ ). We can also describe these conclusions graphically if we take  $n$  points and connect points corresponding to the quantities  $x_i$  and  $x_j$  by an edge iff these quantities  $x_i$  and  $x_j$  can have equal values.

Graphs obtained in this manner for the situations when all sets  $X_i$  are intervals are called *interval graphs*. Not every graph can be thus represented by intervals. There is a rich literature on algorithms for checking whether the given graph is an interval graph, and for designing simple sequences of intervals  $X_1, \dots, X_n$  that lead to a given interval graph; interval graphs have many useful applications (see, e.g., [5]).

If we allow sets  $X_i$  that are not necessarily intervals, then, as one can easily see, every finite graph  $G = (V, E)$  with set of vertices  $V$  and set of edges  $E$  can be represented in this manner: e.g., we can associate with every vertex  $v$  a set  $X_v$  consisting of all the edges that contain  $v$ ; then,  $X_v \cap X_w \neq \emptyset$  iff  $v$  and  $w$  are connected by an edge. This result is known since 1945 [6].

In this proof, we need as many elements in all the sets  $X_v$  as there are edges in the graph; as a result, to represent a graph with  $n$  vertices, we need, in the worst case,  $n(n+1)/2$  elements. A natural question is: do we necessarily need that many elements to represent a given graph?

In [3], Erdős has shown that we can use only half of this number; namely:

- we can always find a representation with  $\lfloor n^2/4 \rfloor$  elements (in  $X_1 \cup \dots \cup X_n$ ), and
- for every integer  $n > 0$ , there exists a graph for which no representation with fewer than  $\lfloor n^2/4 \rfloor$  elements is possible.

### 3. Erdős's results related to (generalized) interval computations: second example

Another result of Erdős is related to a similar problem:

- In the first example, we analyzed the following question: "if the measurement result is  $\bar{x}$ , what are the possible actual values of the measured quantity?". We assumed that for every  $\bar{x}$ , we know the set  $X$  of possible actual values of the measured quantity.

- Due to uncertainty, if we measure the same quantity twice, we may get different results. It is therefore natural to ask a different question: if we know the result  $\tilde{x}$  of a certain measurement, what other values can we get if we measure the same quantity once again? Let us denote this set of values by  $P(\tilde{x})$ .

In this second example, we will assume that we know, for each possible measurement result  $\tilde{x}$  (i.e., for each real number), this set  $P(\tilde{x})$ . It is natural to assume that this set  $P(\tilde{x})$  is bounded (otherwise, if the error can be as large as possible, what good are the measurements?).

We want to ask the same question as in the first example: if we have different measurement results, when can we conclude that these results come from measuring different actual values? If we have two measurement results  $\tilde{x}$  and  $\tilde{x}'$ , and we know that  $\tilde{x}'$  was measured after  $\tilde{x}$ , then the definition of the set  $P(\tilde{x})$  leads to the following simple answer to this question:

- if  $\tilde{x}' \in P(\tilde{x})$ , then  $\tilde{x}'$  and  $\tilde{x}$  can be the results of measuring the same quantity;
- if  $\tilde{x}' \notin P(\tilde{x})$ , then  $\tilde{x}$  and  $\tilde{x}'$  come from measuring different quantities.

If we do not know in what order the measurements were made, then in order to guarantee that  $\tilde{x}$  and  $\tilde{x}'$  come from measuring different quantities we have to assume that both  $\tilde{x}' \notin P(\tilde{x})$  and  $\tilde{x} \notin P(\tilde{x}')$ .

If the set  $P(\tilde{x})$  is simply an interval  $[\tilde{x} - \Delta, \tilde{x} + \Delta]$ , then, of course, the conditions  $\tilde{x} \in P(\tilde{x}')$  and  $\tilde{x}' \in P(\tilde{x})$  mean the same thing: that  $|\tilde{x}' - \tilde{x}| \leq \Delta$ . However, this is not necessarily always so: In real life, it is possible that the properties of a measuring instrument change with time (e.g., there may be a minor drift of its characteristics), and therefore, the relation  $\tilde{x}' \in P(\tilde{x})$  may be asymmetric.

For this situation, Erdős formulated, in [1, 4], the following problem:

- If the relation  $\tilde{x}' \in P(\tilde{x})$  is symmetric, and if all sets  $P(\tilde{x})$  are bounded, then one can easily prove that for every  $n$ , there exist  $n$  real numbers (membership results)  $\tilde{x}_1, \dots, \tilde{x}_n$  that represent  $n$  different actual values, i.e., for which  $\tilde{x}_i \notin P(\tilde{x}_j)$  for all  $i \neq j$ .
- However, for asymmetric relation  $\tilde{x}' \in P(\tilde{x})$ , this is not necessarily true.

When is it true? The answer provided in [1, 4] is: If the (Lebesgue) *measure* of all sets  $P(\tilde{x})$  is bounded by a constant  $M$ , then for every  $n$ , it is possible to have  $n$  measurement results  $\tilde{x}_1, \dots, \tilde{x}_n$  that imply that no two actual values are equal.

## 4. There are probably many more

Erdős's legacy is huge, and probably there are more examples of interval-related results; even papers that are not directly related to verified computing may have influenced (or will influence) other papers that are of importance to us.

## References

- [1] Erdős, P. *Some unsolved problem*. Magyar Tud. Akad. Mat. Kut. Int. Közl., 1961, pp. 221–254; partially reprinted in [2], pp. 15–22.

- [2] Erdős, P. *The art of counting. Selected writings*. Spencer, J. (ed.), MIT Press, Cambridge, MA, 1973.
- [3] Erdős, P., Goodman, A. W., and Pósa, L. *The representation of graph by set intersection*. *Canad. Math. J.* **18** (1966), pp. 106–112; reprinted in [2], pp. 87–93.
- [4] Erdős, P. and Hajnal, A. *Some remarks on set theory. VIII*. *Michigan Math. Journal* **7** (1960), pp. 187–191.
- [5] Fishburn, P. *Interval orders and interval graphs*. Wiley, N.Y., 1985.
- [6] Szpilrain-Marczewski, E. *Sur deux propriétés des classes d'ensembles*. *Fund. Math.* **33** (1945), pp. 303–307.

V. KREINOVICH  
Computer Science Department  
University of Texas at El Paso  
El Paso, TX 79968, USA