EXPLICIT GRAPHS WITH EXTENSION PROPERTIES

ANDREAS BLASS AND BENJAMIN ROSSMAN

Abstract. We exhibit explicit, combinatorially defined graphs satisfying the k^{th} extension axiom: Given any set of k distinct vertices and any partition of it into two pieces, there exists another vertex adjacent to all of the vertices in the first piece and to none in the second.

Quisani:[1] I've been reading about zero-one laws, and many of the results involve *extension axioms*. In the simple case of graphs, by which I mean undirected graphs without loops or multiple edges, the k^{th} extension axiom[2] says that, for any k distinct vertices $x_1, \dots, x_k$ and any subset $S \subseteq \{1, \dots, k\}$, there is another vertex adjacent to x_α for all $\alpha \in S$ and for no other α. I know that each of these axioms is true in almost all sufficiently large finite graphs. (Of course, "sufficiently large" depends on k.) So there are lots of these graphs, but I'd like to see some actual examples.

Authors: Well, just take a big set of vertices, flip coins to decide which pairs to join by edges, and chances are you'll get what you want.

Q: Yes, but I'd like a reasonably regular-looking graph, not something totally random.

A: There are strongly regular graphs that satisfy extension axioms. In fact, Cameron and Stark [3] show that there are lots of them.

Q: What does "strongly regular" mean?

A: Every vertex has the same number of neighbors (i.e., the graph is regular), every pair of adjacent vertices has the same number of common neighbors, and similarly for pairs of non-adjacent vertices.

Q: That sounds good; so they got rid of the randomness.

A: No, part of their construction involves randomization.

Q: So they don't get a really explicit example? That's what I'd want — an example that I can get my hands on and really see why the k-extension axiom holds. Have people given explicit, non-randomizing constructions of graphs that satisfy extension axioms? More precisely, are there explicitly defined sequences of finite graphs such that, as you

[1]We thank Yuri Gurevich for lending us his graduate student, Quisani.

[2]Extension axioms are also called *adjacency axioms*, and graphs that satisfy the k^{th} extension axiom are also called *k-existentially closed*.

go further in the sequence, more and more of the extension axioms are true?
A: Yes. It was shown in both [1] and [2] that Paley graphs have the property you want.[3] Furthermore, they are strongly regular.
Q: What are Paley graphs?
A: Take a prime p (or a prime power) that is congruent to 1 modulo 4 and form a graph whose vertices are the elements of the field $\mathbb{F}_p$ of size p. Join two distinct vertices by an edge if their difference is a square in $\mathbb{F}_p$. The theorem is that this graph will satisfy the k^{th} extension axiom provided p is sufficiently large compared to k.
Q: Three questions: How large is sufficiently large? Can you explain why these Paley graphs satisfy extension axioms? And what's the purpose of having $p \equiv 1 \pmod 4$?
A: The easiest question to answer is the last. The congruence is needed to ensure that -1 is a square in $\mathbb{F}_p$, which in turn is needed to make sure the Paley graph is an undirected graph; x is joined to y if and only if y is joined to x. If p were $\equiv 3 \pmod 4$ then we'd have a Paley tournament instead of a graph.

Sufficiently large is exponentially larger than k, roughly $k^2 4^k$. (For comparison, a random graph has a good chance of satisfying the k^{th} extension axiom when the number of vertices is somewhat larger than $k^2 2^k$; see [4].)

Unfortunately, it's not easy to explain why the Paley graphs satisfy the extension axioms. Both [1] and [2] invoke non-trivial results from number theory in the proofs of the extension axioms.
Q: So are there no explicit examples where one can directly see why the extension axioms hold?
A: Actually, we have such examples.[4] They're not as pretty as the Paley graphs, and they're larger (for a given k), but we can explain what they are and how they work, without appealing to any deep theorems.
Q: Great! Show me.
A: OK. Given k, we'll construct a graph whose vertices are certain matrices of 0's and 1's. These matrices will have $r = 2k(k-1)+1$ rows

[3]The result isn't explicitly stated in [2], but it's proved in the course of proving Theorem 3, which says that every finite graph occurs as an induced subgraph in all sufficiently large Paley graphs.

[4]The first version of these examples was derived by the second author from a construction, introduced for quite different purposes, in [8].

and c columns, where c is chosen large enough so that

$$2^c \geq 2^{k^2}\binom{rc}{k-1}$$

Q: Wait a minute; let me check that such a c exists. Yes. The left side is exponential in c while the right side, despite the exponential dependence on k, is only a polynomial in c when k is fixed, so any sufficiently large c will do.
A: Right. Having fixed suitable r and c, let the vertices of our graph be r by c matrices of 0's and 1's in which a majority of the r rows are identical. That is, in each of our matrices, at least $k(k-1)+1$ of the rows are identical.

To define the edges of our graph requires some preliminary terminology. We consider *constraints*, which a vertex may or may not satisfy. A constraint[5] is given by a pair (A, F) where A is a set of $k-1$ locations in our matrix (i.e., $k-1$ pairs (i,j) with $1 \leq i \leq r$ and $1 \leq j \leq c$) and F is a family of at most k functions from A to $\{0,1\}$. We say that a vertex V *satisfies* a constraint (A, F) if the entries in V at the locations in A form an element of F, i.e., if

$$(\exists f \in F)(\forall (i,j) \in A)\, V_{ij} = f(i,j).$$

We need to estimate the number of constraints (A, F). There are $\binom{rc}{k-1}$ possibilities for A. For each fixed A, there are 2^{k-1} functions from A to $\{0,1\}$, so there are $2^{(k-1)k}$ sequences of k such functions. Every possible second component F of a constraint (A, F), except for $F = \varnothing$, is the range of such a sequence. So the number of F's for a fixed A is certainly at most 2^{k^2}, and the total number of constraints is no more than

$$\binom{rc}{k-1}2^{k^2} \leq 2^c.$$

Therefore, we can fix a function C from the set of c-component vectors of 0's and 1's onto the set of constraints. Since the notion "c-component vector of 0's and 1's" will be needed repeatedly, we abbreviated it as "row vector," which makes sense since these are the vectors that occur as rows in our matrices.
Q: You're not choosing C at random, are you?
A: No. We promised an explicit construction, with no randomization. To get a definite C, list all the row vectors in lexicographic order, and, after choosing some reasonable notation for constraints, list the

[5]Readers familiar with combinatorial set theory will notice a similarity between the notion of constraint and Hausdorff's construction [7] of large independent families of sets.

constraints lexicographically also. Then let C map the n^{th} element of the first list to the n^{th} element of the second list, cycling back to the beginning of the second list if the first list is longer (which in fact it will be).

Q: OK. It's an unpleasantly arbitrary C, but I agree it's not random. Why do you cycle back to the beginning rather than, say, just repeating the last element?

A: The cycling is irrelevant in this argument, but we'll want it for another purpose later.

You're quite right about the arbitrariness of C. Any surjection C will work for this proof, so, if you can think of a nicer explicit C, feel free to use it. But remember, we warned you that these graphs won't be as pretty as Paley graphs.

Using C, every vertex V of our graph determines a constraint V^* as follows. A majority of the rows of V are the same row vector, which we call the *majority row* of V; apply C to that vector to get a constraint V^*.

Now define a directed graph by putting an arrow from V to W whenever W satisfies the constraint V^*.

Q: I thought you were going to produce an *undirected* graph.

A: We will; the directed graph is only an auxiliary construction. The undirected graph has an edge joining V and W if, of the two possible directed edges, V to W and W to V, either both are present or neither is present. That is, V is adjacent to W just in case

$$(V \text{ satisfies } W^*) \iff (W \text{ satisfies } V^*).$$

We'll show that the graph so defined satisfies the k^{th} extension axiom, but in order to do so we'll need the following preliminary information.

Claim. Let $V_1, \dots, V_k$ be k distinct vertices of our graph, and let S be any subset of $\{1, \dots, k\}$. There is a constraint that is satisfied by V_α for all $\alpha \in S$ and for none of the other α's.

Proof. It suffices to find a set A of $k-1$ locations (i, j) that separate the V_α's, in the sense that, whenever α and β are distinct indices in $\{1, \dots, k\}$, then V_α and V_β differ at some location in A. Once we have such an A, we have k distinct functions $f_\alpha : A \to \{0, 1\}$ defined by $f_\alpha(i, j) = (V_\alpha)_{ij}$. Then let $F = \{f_\alpha : \alpha \in S\}$ and observe that (A, F) can serve as the desired constraint. So it remains only to produce an appropriate A.

Q: That would be trivial if you allowed $\binom{k}{2}$ locations in A, rather than only $k-1$. You could just choose, for each V_α and V_β, one location where they differ.

A: Right, and in fact, if you don't want to worry about getting $|A|$ down to $k-1$, you could rewrite this whole story with $\binom{k}{2}$ in place of $k-1$, starting with the definition of c (where the exponent k^2 would also have to be adjusted).

But in fact, it's not hard to achieve $|A| = k-1$. Proceed by induction on k, the case $k = 1$ being vacuous. For $k > 1$, start by choosing a location (i, j) where some V_α and V_β differ. So our set of k vertices is partitioned into two nonempty subsets, according to the (i, j) entries. Let these subsets consist of a and b elements, so $a+b = k$. By induction hypothesis, we can find $a-1$ locations sufficient to separate any two vertices from the first class and $b-1$ locations sufficient to separate any two vertices from the second class. Together with (i, j), that gives us $(a-1)+(b-1)+1 = k-1$ locations that separate all the vertices. $\square$

Q: The claim you just proved gives a sort of extension axiom for the auxiliary, directed graph. The constraint from the claim is $C(w)$ for some row vector w. An $r \times c$ matrix W having all its rows equal to w would be a vertex of your graph, and there would be a directed edge from W to V_α if and only if $\alpha \in S$. So if you could arrange for this W to satisfy all the constraints V_α^*, then W would be adjacent, in the undirected graph, to V_α if and only if $\alpha \in S$.

Unfortunately, I don't see how you can arrange that. The constraints V_α^* might contradict each other.

A: That's right, so we have to be a little sneakier.

Suppose we're given distinct vertices $V_1, \dots, V_k$ and a set $S \subseteq \{1, \dots, k\}$ as above, and we want a vertex W adjacent to V_α if and only if $\alpha \in S$.

First, fix an arbitrary (not random!) vertex W'; for definiteness, let it be the matrix of all zeros. Let

$$T = \{\alpha \in \{1, \dots, k\} : W' \text{ satisfies } V_\alpha^*\}.$$

Apply the claim with the given vertices V_α but with S replaced by the complement of the symmetric difference of S and T, i.e., by $\{\alpha : (\alpha \in S) \iff (\alpha \in T)\}$. The constraint given by the claim is, as you noted, $C(w)$ for some row vector w. Let W'' be the vertex that has all its rows equal to w. So we have, thanks to the choice of W'' and the definition of T,

$$V_\alpha \text{ satisfies } (W'')^* \iff \big((\alpha \in S) \iff (W' \text{ satisfies } V_\alpha^*)\big).$$

Since $\iff$ is an associative and commutative operation on truth values, this can be rewritten as

$$\alpha \in S \iff \big((V_\alpha \text{ satisfies } (W'')^*) \iff (W' \text{ satisfies } V_\alpha^*)\big). \tag{1}$$

Q: You'd be done if W' and W'' were equal, but that would require a miracle. It's true that W' was arbitrary, but W'' depends on T which depends on the choice of W', so I see no chance to use the arbitrariness of W' to make it match W''.
A: Absolutely right; there's no reason to think W' and W'' are equal. But we can combine them into a single W that inherits the desirable features of both.

For each α, whether W' satisfies the constraint V^*_α (and thus whether $\alpha \in T$) depends only on the entries of the matrix W' in $k-1$ locations, namely the locations in the first component A of the constraint $V^*_\alpha = (A, F)$. So at most $k(k-1)$ entries of W' are involved in the satisfaction or non-satisfaction of the k constraints V^*_α. Define W to agree with W' in those entries and with W'' at all other locations.

We've kept enough entries of W' in W to ensure that

$$(W \text{ satisfies } V^*_\alpha) \iff (W' \text{ satisfies } V^*_\alpha). \tag{2}$$

On the other hand, W agrees with W'' at all but at most $k(k-1)$ entries. Since there are $r = 2k(k-1)+1$ rows, the majority of the rows of W are identical to the rows of W'', namely the w that we chose when constructing W''. So

$$W^* = (W'')^*. \tag{3}$$

Inserting (2) and (3) into (1), we get

$$\alpha \in S \iff \big((V_\alpha \text{ satisfies } W^*) \iff (W \text{ satisfies } V^*_\alpha)\big).$$

So W is as required by the extension axiom.
Q: That's a clever proof. You could have gotten by with a slightly smaller graph, if you had been less generous when estimating the number of constraints. The factor 2^{k^2} is larger, for any $k > 1$, than the value actually given by your argument, $2^{k(k-1)} + 1$. Even that can be reduced since a typical family F will have many sequences enumerating it in different orders, and with different repetitions in case $|F| < k$.
A: That's right. And other reductions are possible. For example, we could define the vertices of our graph to be only those r by c matrices in which all the rows are identical except for at most $k(k-1)$ locations where 1's have been changed to 0's. These vertices suffice, because they include every W used in the proof of the k-extension property (since W'' had all its rows identical and we took W' to be the all 0 matrix).

Another substantial reduction could be obtained by being more clever in our choice of error-correcting code.
Q: I didn't see any error-correcting code here.

A: When we took a row vector w and repeated it r times to make a matrix W'', we were in effect using the simplest error-correcting code, namely repetition. The point of the repetition is that when we changed at most $k(k-1)$ entries of W'' to form W, the original w could still be recovered, despite the changes. That's exactly what error-correcting codes are good for. By using a more sophisticated code, we could get by with vertices that contain far fewer than rc binary components, and so we could get a smaller graph.

Furthermore, since the "errors" that we introduced into a "code word" W'' to produce our W were only replacing some 1's by 0's, never the reverse, the code only has to correct errors of this one sort.

Q: This improvement looks pretty complicated, especially since I know almost nothing about coding theory. Rather than going into the details, it might be more interesting to look for other extension properties that can be obtained by the same method.

A: A slight modification of the method gives tournaments satisfying the natural extension axioms.

Q: Presumably, the k^{th} of these natural extension axioms for tournaments says that, given k distinct vertices $V_1, \dots, V_k$ and given a subset $S \subseteq \{1, \dots, k\}$, there is another vertex W with a directed edge to V_α when $\alpha \in S$ and a directed edge from V_α when $\alpha \notin S$.

A: That's right. The same arguments as for undirected graphs show that each of these extension axioms holds in almost all sufficiently large finite graphs and that the extension axioms plus the basic axioms for tournaments (saying that, for each pair of distinct vertices, there is an edge between them in exactly one direction and that there are no loops) constitute a complete first-order theory.[6] Thus, one gets a zero-one law for the first-order[7] properties of tournaments.

Also, recall that the Paley construction with $p \equiv 3 \pmod 4$ produces tournaments rather than undirected graphs. It is known [6] that these Paley tournaments satisfy the k^{th} extension axiom[8] provided p is large enough compared to k. As with the Paley graphs, any $p \geq k^2 4^k$ is large enough, and the proof relies on the same non-trivial number theory.

Q: I suppose that your construction of undirected graphs can be converted into an analogous construction for tournaments by replacing

[6] As in the undirected case, one can prove completeness either by showing that the theory admits elimination of quantifiers or by showing that it is categorical in power $\aleph_0$.

[7] This extends easily to finite-variable infinitary logic.

[8] The result stated in [6] is weaker, namely that, given any k vertices, there is another vertex with edges directed to each of the given ones. A minor modification of the proof, however, would establish the extension axiom.

"adjacent" and "non-adjacent" with "edge directed one way" and "edge directed the other way," right?

A: That's the basic idea, but we need to clarify what is "one way" and what is "the other way."

Q: Can't you just linearly order the vertices and let "one way" mean from the earlier to the later vertex in this ordering?

A: That doesn't quite work. The problem is that, if you're given the vertices V_α and the set S specifying the directions for the edges between the V_α's and the desired W, you need to convert these data into specifications of adjacency or non-adjacency between the V_α's and W in the undirected graph. These new specifications will depend on the position of W relative to the V_α's in your ordering. By Murphy's law, if you choose a particular relative position for W, you'll get constraints that can only be satisfied by W's in other relative positions, not the one you chose.

Q: I understand the problem. How do you escape from it?

A: Here's a modification of the undirected construction. First, increase c if necessary so that

$$2^c \geq (k+1)2^{k^2}\binom{rc}{k-1}.$$

This ensures that, when we produce the map C from row vectors to constraints, each constraint is the image of at least $k+1$ row vectors. In fact, because of the specific way we obtained C (which was irrelevant earlier but is important now), as w runs through all the row vectors in lexicographic order, $C(w)$ will cycle through the contraints at least $k+1$ times. Thus, C has the following "interval property":

- The lexicographic order of the row vectors contains $k+1$ disjoint intervals, each of which is mapped by C onto all of the constraints.

Second, list the vertices in such an order that, if the majority row of V lexicographically precedes the majority row of W then V precedes W in the list.

Q: That implies that all the vertices with a given majority row occur consecutively in the list.

A: Right, and that will be important for our proof.

The interval property above and the specification of our list have the following consequence, an "interval property" for vertices:

- The list of vertices contains $k+1$ disjoint intervals such that, if we are given one of these intervals and we are given a constraint, then there exists a row w such that $C(w)$ is the given constraint and all vertices with majority row w are in the given interval.

Now define a tournament just as you suggested above: There is an edge from V to W if and only if either V and W are adjacent in our undirected graph and V precedes W in our list or V and W are not adjacent in the undirected graph and W precedes V in the list.

To prove that this tournament satisfies the k^{th} extension axiom, let distinct vertices $V_1, \dots, V_k$ be given, along with a set $S \subseteq \{1, \dots, k\}$. Since there are only k given vertices V_α and there are $k+1$ intervals in the interval property for vertices, fix one of these intervals, say $\mathcal{I}$, that contains no V_α. Thus, the set

$$U = \{\alpha : W \text{ precedes } V_\alpha \text{ in the list}\}$$

is the same for all vertices $W \in \mathcal{I}$.

In our proof of the k^{th} extension axiom for our undirected graph, we found a row w coding a certain constraint and then we constructed the required W to have majority row w (while a minority of rows provided sufficient agreement with W'). In the present context, we can always, thanks to the interval property, choose w so that all vertices W with majority row w are in $\mathcal{I}$. We use this to find some $W \in \mathcal{I}$ that is adjacent, in the undirected graph, to V_α exactly when α is in both or neither of S and U, that is, when

$$(\alpha \in S) \iff (\alpha \in U).$$

Then, for each $\alpha \in S$, we have that each of the following statements is equivalent to the next.

- There is a directed edge from W to V_α.
- (V_α and W are adjacent in the undirected graph) $\iff$ (W precedes V_α in the list).
- $\big((\alpha \in S) \iff (\alpha \in U)\big) \iff (\alpha \in U)$.
- $\alpha \in S$.

Q: So you got around the "Murphy's law" problem by making sure that you could specify the position of W relative to the V_α's and still have enough W's to obtain the desired adjacencies and non-adjacencies.

Can your construction be used to get extension axioms in more contexts?

A: Probably, but we haven't yet looked into this carefully. It would be particularly interesting to get triangle-free graphs satisfying the appropriate extension axioms. These axioms are the same as for ordinary, undirected graphs, except that they assume the vertices V_α for $\alpha \in S$ are pairwise non-adjacent. That's obviously necessary so that the vertex W given by the axiom doesn't complete a triangle. What makes this case particularly interesting is that it is not known that the desired objects — finite triangle free graphs satisfying the k^{th} extension axiom for a prescribed k — exist at all. The probabilistic arguments used to give existence proofs in the case of graphs and tournaments do not apply to the case of triangle-free graphs. Nor is there a Paley-style construction; indeed it is known that a strongly regular triangle-free graph cannot satisfy the 4^{th} extension axiom (see [5]).

References

[1] Andreas Blass, Geoffrey Exoo, and Frank Harary, "Paley graphs satisfy all first-order adjacency axioms," *J. Graph Theory* 5 (1981) 435–439.

[2] Béla Bollobás and Andrew G. Thomason, "Graphs which contain all small graphs," *European J. Combin.* 2 (1981)13–15.

[3] Peter J. Cameron and Dudley Stark, "A prolific construction of strongly regular graphs with the n-e.c. property," *Electron. J. Combin.* 9 (2002) Research paper 31 (12 pages).

[4] Louis Caccetta, Paul Erdős, and Kaipillil Vijayan, "A property of random graphs," *Ars Combin.* 19 (1985) 287–294.

[5] Gregory Cherlin, "A problem on triangle-free graphs," http://www.math.rutgers.edu/∼cherlin/Problems/tfree.html

[6] Ronald L. Graham and Joel H. Spencer, "A constructive solution to a tournament problem," *Canad. Math. Bull.* 14 (1971) 45–48.

[7] Felix Hausdorff, "Über zwei Sätze von G. Fichtenholz und L. Kantorovitch," *Studia Math.* 6 (1936) 18–19.

[8] Benjamin Rossman, "Successor-Invariance in the Finite," *Proceedings of the 18th IEEE Symposium on Logic in Computer Science* (2003) 148–157.

Mathematics Department, University of Michigan, Ann Arbor, MI 48109–1109, U.S.A.

E-mail address: ablass@umich.edu

Microsoft Research, One Microsoft Way, Redmond, WA 98052, U.S.A.

E-mail address: brossman@impa.br