Buckets Graph Description

(in progress — *warning* some information might be incorrect or incomplete)

This is meant to be fast in implementation. The buckets graph has degree $2m$ and there are m times as many buckets as things in buckets, and two things match if their buckets are connected so there's no need for a further comparison function. Because many nearby things are compared to buckets at the exact same offset, it's possible to implement this efficiently by making a bitfield of which buckets have something in them then doing bitshifts and & and comparing to zero.

The buckets graph is organized as follows:

Buckets are grouped into b-groups, and b-groups are grouped into c-groups. "B-group size" refers to the number of buckets per b-group, and "c-group size" refers to the number of b-groups per c-group.

Let $BgrSz$ be the b-group size, let $CgrSz$ be the c-group size, and let $numCgr$ be the total number of c-groups.

Then the number of b-groups is $numBgr = CgrSz * numCgr$, and the number of buckets is $numBuc = BgrSz * numBgr$.

Let the total degree of each bucket be $2m$ for $m \ni \mathbb{Z}$.

The four variable parameters for graph construction, then, are: $BgrSz$, $CgrSz$, $numCgr$, and m.

For each edge of the graph, we define 3 offsets: the c-group offset, the b-group offset, and the bucket offset. The c-group offset is 1 for all edges. The b-group offset (which is the offset between b-groups within a c-group) is r, where r ranges as $0 \le r < m$. And the bucket offset (which is the offset between buckets within a b-group) is q, where for outgoing edges from buckets in even-indexed c-groups, $q = (2r)^2$, $0 \le r < m$, and for outgoing edges from buckets in odd-indexed c-groups, $q = (2r+1)^2$, $0 \le r < m$. The bucket offsets therefore alternate between $q = (2r)^2$ and $q = (2r+1)^2$ from c-group to c-group.

For each bucket x, let $indI$ be x's c-group index, let $indJ$ be x's b-group index within the c-group, and let $indK$ be x's bucket index within the b-group. We can define these as follows:

$$indI = floor\left\{\frac{x}{BgrSz*CgrSz}\right\}, \qquad indJ = floor\left\{\frac{x-(indI*BgrSz*CgrSz)}{BgrSz}\right\},$$
$$indK = x - [(indI * BgrSz * CgrSz) + (indJ * BgrSz)].$$

The set of buckets $\{y_i\}$ connected to x via x's outgoing edges is given by:
$y_r = \{[(indI + 1) \% numCgr] * BgrSz * CgrSz\} + \{[(indJ + r) \% CgrSz] * BgrSz\} + \left[(q^2 + x) \% BgrSz\right]$, for each r in the range $0 \le r < m$, and such that if x is located in an even-indexed c-group, then $q = (2r)^2$, whereas if x is located in an odd-index c-group, then $q = (2r+1)^2$.

The set of buckets connected to x via x's incoming edges are all of those buckets whose set of outgoing connections $\{y_i\}$ is such that $x \ni \{y_i\}$.

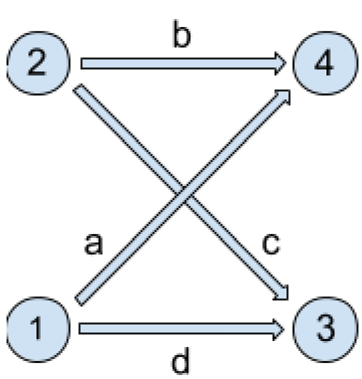
For the bucket graph used in Chik's proof-of-space, each bucket has $deg = 2m = 64$, with $m = 32$ outgoing edges and $m = 32$ incoming edges.

The nodes graph is inherited from the buckets graph. A comparison function is unnecessary because buckets only rarely contain a node and much more rarely contain multiple nodes.

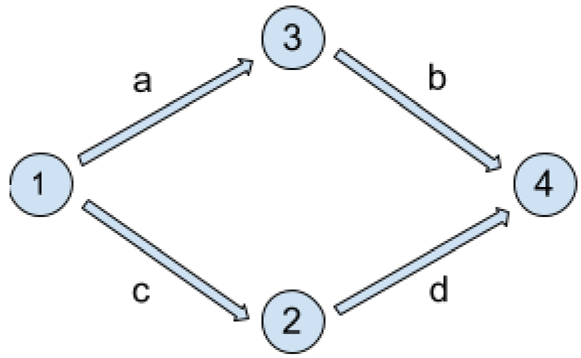
Let us now prove that this graph contains no 4-cycles. [[Need to add something here about the parameter constraints under which the graph contains no 4-cycles.]]

Consider the following two types of 4-cycles:

Type 4-A:



Type 4-B:

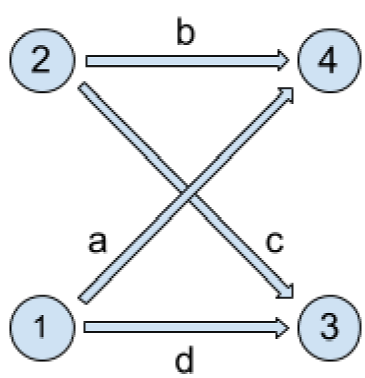


We first consider the following construction, different from the one described above, in which for each bucket, and for r ranging as $0 \leq r < m$, we have the following: the c-group offset is 1, the b-group offset within the c-group is r, and the bucket offset within the b-group is r^2. The set of buckets $\{y_i\}$ connected to x via x's outgoing edges is therefore given by:

$$y_r = \{[(indI + 1)\,\%\,numCgr] * BgrSz * CgrSz\} + \{[(indJ + r)\,\%\,CgrSz] * BgrSz\} + \left[(x + r^2)\,\%\,BgrSz\right]$$

, for each r in the range $0 \leq r < m$.

Type 4-A:



Such a 4-cycle would occur when both $a - b + c - d = 0 \;\; mod\ CgrSz$ and $a^2 - b^2 + c^2 - d^2 = 0 \;\; mod\ BgrSz$. For now we will disregard the different moduli and address that issue later on.

Assuming that $a - b + c - d = 0$, we want to check for the conditions under which it's possible that $a^2 - b^2 + c^2 - d^2 = 0$.

Note the identity $a - b = d - c$, which can be rearranged to express each variable in terms of the other three.

We then have:

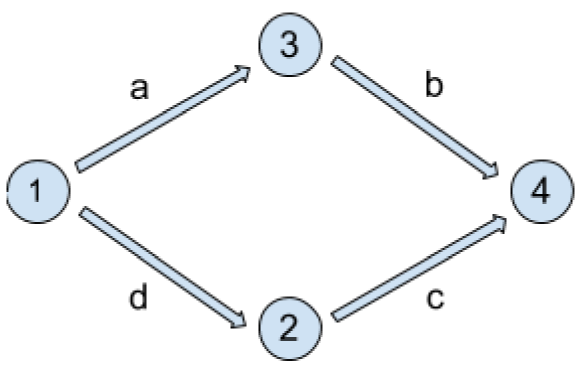
$$a^2 - b^2 + c^2 - d^2 = a^2 - b^2 + c^2 - (a - b + c)^2 = a^2 - b^2 + c^2 - (a^2 - 2ab + 2ac + b^2 - 2bc + c^2)$$
$$= 2ab - 2ac - 2b^2 + 2bc = 2(ab - ac - b^2 + bc)$$

Setting this expression equal to zero, we have:

$$2(ab - ac - b^2 + bc) = 0 \Rightarrow ab - ac - b^2 + bc = 0 \Rightarrow (a - b)(b - c) = 0$$

Therefore, a 4-cycle of type 4-A is possible whenever $a = b$ or $b = c$. Equivalently, carrying out this prescription but instead substituting for the other variables — c, b, and a — yields (respectively) the following conditions: $a = b$ or $a = d$; $a = d$ or $c = d$; and $b = c$ or $c = d$. Clearly, 4-cycles of this kind appear only in cases when an edge is identical to the previous edge in a sequence, indicating a repeated edge. Because our graph does not allow repeated edges, 4-A cycles are not possible in this graph.

Type 4-B:



Such a 4-cycle would occur when both $a + b - c - d = 0 \ mod\ CgrSz$ and $a^2 + b^2 - c^2 - d^2 = 0 \ mod\ BgrSz$. For now we will disregard the different moduli and address that issue later on.

Assuming that $a + b - c - d = 0$, we want to check for the conditions under which it's possible that $a^2 + b^2 - c^2 - d^2 = 0$.

Note the identity $a + b = c + d$ which can be rearranged to express each variable in terms of the other three.

We have:

$$a^2 + b^2 - c^2 - d^2 = a^2 + b^2 - c^2 - (a + b - c)^2 = a^2 + b^2 - c^2 - (a^2 + 2ab - 2ac + b^2 - 2bc + c^2)$$
$$= -2ab + 2ac - 2bc - 2c^2 = -2(ab - ac + bc + c^2)$$

Setting this expression equal to zero, we have:

$$-2(ab - ac + bc + c^2) = 0 \Rightarrow ab - ac + bc + c^2 = 0 \Rightarrow (a - c)(b - c) = 0$$

Therefore, a 4-cycle of type 4-B is possible whenever $a = c$ or $b = c$. Equivalently, carrying out this prescription but instead substituting for the other variables — c, b, and a — yields (respectively) the following conditions: $a = d$ or $b = d$; $a = c$ or $a = d$; and $b = c$ or $b = d$. It is obvious that the conditions in which equal edges share a node (i.e. $b = c$ and $a = d$) are impossible because our graph does not allow repeated edges. However, the conditions in which the equal edges do not share a node do allow 4-cycles to occur. We address this issue in the following way.

Note that the conditions which concern us are $a = c$ and $b = d$. Say that bucket 1, from which edges a and d originate, is located in the c-group at index i. Then buckets 2 and 3, from which edges c and b, respectively, originate, are located in the c-group at index $i + 1$. We therefore impose the following rule: for a bucket in an even-indexed c-group, let its outgoing edges have bucket offsets r^2 ranging through each even-valued r in the range $0 \le r < m$, and for a bucket in an odd-indexed c-group, let its outgoing edges have bucket offsets r^2 for each odd-valued r in the range $0 \le r < m$. The bucket offsets r^2 therefore alternate from c-group to c-group between using even-valued r values and using odd-valued r values.