RECONSTRUCTING COMPOSITIONS

VINCENT VATTER*
School of Mathematics and Statistics
University of St Andrews
St Andrews, Fife, Scotland
vince@mcs.st-and.ac.uk
http://turnbull.mcs.st-and.ac.uk/~vince

We consider the problem of reconstructing compositions of an integer from their subcompositions, which was raised by Raykova (albeit disguised as a question about layered permutations). We show that every composition w of $n \geq 3k+1$ can be reconstructed from its set of k-deletions, i.e., the set of all compositions of $n-k$ contained in w. As there are compositions of $3k$ with the same set of k-deletions, this result is best possible.

Introduction. The Reconstruction Conjecture states that given the multiset of isomorphism types of 1-vertex deletions (briefly, 1-*deletions*) of a graph G — the *deck* of G — on three or more vertices, it is possible to determine G up to isomorphism. The stronger set version of the conjecture due to Harary [5] only allows access to the *set* of 1-deletions and requires G to have four or more vertices. These conjectures can be made even more difficult by considering k-deletions instead of 1-deletions, for which we refer to Manvel [7].

Such reconstruction questions extend naturally to other combinatorial contexts. For example, Schützenberger and Simon (see Lothaire [6, Theorem 6.2.16]) proved that every word of length $n \geq 2k+1$ can be reconstructed from its set of k-deletions (i.e., subwords of length $n-k$). This bound is tight because the words $(ab)^k$ (the word with ab repeated k times) and $(ba)^k$ have the same set of k-deletions: all words of length k over the alphabet $\{a, b\}$. Answering a question of Cameron [4], Pretzel and Siemons [8] considered the partition context, where they proved that every partition of $n \geq 2(k+3)(k+1)$ can be reconstructed from its set of k-deletions.

Motivated by a question of Raykova [9] (described at the end of the paper), we consider the problem of set reconstruction for compositions (ordered partitions), establishing the following result.

---

*Supported by EPSRC grant GR/S53503/01.
Date: March 20, 2007
Key words and phrases. composition, layered permutation, reconstruction
AMS 2000 Subject Classification. 68R15, 06A07, 05A17, 05C60

Theorem 1. *All compositions of $n \geq 3k+1$ can be reconstructed from their sets of k-deletions.*

Our proof of Theorem 1 illustrates an algorithm to perform the reconstruction. Perhaps more convincing than the proof is the Maple implementation of this algorithm, available from the author's homepage.

Notation. We view a composition as a word w whose letters are positive integers, i.e., a word in $\mathbb{P}^*$. We denote the length of w by $|w|$ and the sum of the entries of w by $\|w\|$, and say that w is a composition of $\|w\|$. A 1-deletion of w is a composition that can be obtained either by lowering a ≥ 2 entry of w by 1 or by removing an entry of w that is equal to 1. A 2-deletion is then a 1-deletion of a 1-deletion, and so on.

This notion naturally defines a partial order† on compositions: $u \leq w$ if w contains a subword $w(i_1)w(i_2)\cdots w(i_\ell)$ of length $\ell = |u|$ such that $u(j) \leq w(i_j)$ for all $1 \leq j \leq \ell$. (We refer to the indices $i_1 < \cdots < i_\ell$ as an *embedding* of u.) For example, $1211 \leq 21312$ because of the subword 2312. If $u \leq w$ then u is a $(\|w\| - \|u\|)$-deletion of w. Returning to the previous example, $\|21312\| = 9$ and $\|1211\| = 5$, so 1211 is a 4-deletion of 21312.

A lower bound. In the context of words, the fact that the sets of k-deletions of $(ab)^k$ and $(ba)^k$ are both equal to the set of all words of length k over $\{a, b\}$ provides a lower bound on k-reconstructibility. Here we can use a very similar example: the sets of k-deletions of $(12)^k$ and $(21)^k$ are both equal to the set of all compositions of $2k$ in which no entry is greater than 2. This implies that Theorem 1 is best possible.

The proof. Our reconstruction algorithm/proof of Theorem 1 employs several composition statistics. One is the *exceedance number*, defined by $\mathrm{ex}(w) = \|w\| - |w| = \sum(w(i) - 1)$ where the sum is over all entries $w(i)$. Another is the number of ones in w, which can be approximated from its set of k-deletions:

Lemma 2. *The composition w of $n \geq 3k+1$ has at least k ones if and only if either*

(1) 1^{n-k} is a k-deletion of w, or

(2) the longest k-deletion of w is k letters longer than the shortest k-deletion of w.

Moreover, w has precisely k ones if and only if (2) holds and w has a k-deletion without ones.

Proof. It is easy to see that if either (1) or (2) occurs then w has at least k ones. Suppose then that w has at least k ones. If $\mathrm{ex}(w) \leq k$ then since $1^{|w|}$ is an $\mathrm{ex}(w)$-deletion of w, it follows that 1^{n-k} is a k-deletion of w, satisfying (1). On the other hand, if $\mathrm{ex}(w) > k$ then some k-deletion of w has length $|w|$, while the fact that w contains at least k ones guarantees that some k-deletion of w has length $|w| - k$, satisfying (2). The second claim in the lemma is then readily verified. □

†This partial order was first considered by Bergeron, Bousquet-Mélou, and Dulucq [1], and has since been studied by Snellman [12, 13], Sagan and Vatter [10], and Björner and Sagan [2].

Given a set of k-deletions of a composition, the first step in our algorithm is to apply Lemma 2 to decide if the composition has fewer than k, precisely k, or more than k ones. The three cases are handled separately. The first two are relatively straightforward, while the last is more delicate.

Lemma 3. *If w is a composition of $n \geq 3k+1$ with fewer than k ones, then w can be reconstructed from its set of k-deletions.*

Proof. Given the set of k-deletions of a composition w satisfying these hypotheses, our algorithm can apply the result of Lemma 2 to determine that w has fewer than k ones. It then follows that

$$\mathrm{ex}(w) \geq \frac{\|w\| - (\text{\# of ones in } w)}{2} \geq \frac{2k+2}{2} = k+1.$$

From this we see that w has the same length, say m, as its longest k-deletions, and thus $\mathrm{ex}(w)$ can be easily determined: it is k plus the exceedance number of one of the longest k-deletions.

Set $t = \mathrm{ex}(w) - k$ and define the composition $a = a(1)\cdots a(m)$ by

$$a(i) = \max\{s : \underbrace{1\cdots1}_{i-1}\, s\, \underbrace{1\cdots1}_{m-i} \text{ is, or is contained in, a } k\text{-deletion of } w\}.$$

It follows that a satisfies

$$a(i) = \min\{w(i), t+1\}. \tag{1}$$

There are now two cases in which we are done:

- If $\|a\| = n$ then w must be equal to a. By (1), this will occur if w contains no entries greater than $t+1$.
- If at most one entry of a satisfies $a(i) = t+1$, which by (1) will occur if w contains at most one entry $w(i) \geq t+1$, then (1) forces $w(j) = a(j)$ for all $j \neq i$ and then $w(i)$ can be calculated from the fact that $\|w\| = n$.

Suppose, for the sake of contradiction, that neither of these conditions hold. Thus w must contain an entry $w(i) > t+1$ and another entry $w(j) \geq t+1$. We then have

$$k + t = \mathrm{ex}(w) \geq t + (t+1) + (\text{\# of} \geq\!2 \text{ entries in } w\text{, not including } w(i), w(j)),$$

so

$$k \geq t + 1 + (\text{\# of} \geq\!2 \text{ entries in } w\text{, not including } w(i), w(j)), \tag{2}$$

while

$$|w| = 2 + (\text{\#1s in } w) + (\text{\# of} \geq\!2 \text{ entries in } w\text{, not including } w(i), w(j)),$$

so because w contains fewer than k ones,

$$(\text{\# of} \geq\!2 \text{ entries in } w\text{, not including } w(i), w(j)) \geq |w| - k - 1. \tag{3}$$

Combining (2) and (3) shows that $|w| \leq 2k-t$, but then $\mathrm{ex}(w) \geq (3k+1)-(2k-t) = k+t+1$, contradicting the definition of t and completing the proof. □

Example 4. Suppose the reconstruction algorithm is given the set of 3-deletions

$$\{52, 322, 412, 421, 511, 2122, 3112, 3121, 4111\}$$

of an unknown composition w of $n = 10$. The algorithm first checks the hypotheses of Lemma 2. The first condition does not hold because the set of 3-deletions does not contain $1^{10-3} = 1111111$, while the second condition fails because the longest 3-deletion is only 2 letters longer than the shortest. Therefore w has fewer than $k = 3$ ones. Now the algorithm follows the proof of Lemma 3. First we compute $\mathrm{ex}(w)$ from one of the longest 3-deletions:

$$\mathrm{ex}(w) = \mathrm{ex}(3121) + 3 = 6,$$

so $t = 3$. Then we compute a:

$a(1) = 4$ because 4111 is contained in a 3-deletion but 5111 is not,
$a(2) = 1$ because 1111 is contained in a 3-deletion but 1211 is not,
$a(3) = 2$ because 1121 is contained in a 3-deletion but 1131 is not,
$a(4) = 2$ because 1112 is contained in a 3-deletion but 1113 is not.

Thus $w \geq 4122$. Since $\|4122\| = 9 < 10 = \|w\|$, we are not done reconstructing w and need to account for one more exceedance. However, since $a(1)$ is the only entry of a equal to $t + 1 = 4$, $w(1)$ is the only entry of w that can be greater than the corresponding entry of a, so we get $w = 5122$.

Lemma 5. *If w is a composition of $n \geq 3k + 1$ with precisely k ones, then w can be reconstructed from its set of k-deletions.*

Proof. Given the set of k-deletions of a composition w satisfying these hypotheses, our algorithm can apply the result of Lemma 2 to determine that it has exactly k ones. With this established, the length of w, say m, can be computed as k plus the length of the shortest k-deletion of w.

There is a k-deletion of w without ones, and this composition gives the ≥2 entries of w in their correct order. Thus it suffices to determine where they lie in w. To this end define the composition a_i by

$$a_i = \underbrace{1 \cdots 1}_{i-1} 2 \underbrace{1 \cdots 1}_{m-i}.$$

If a_i is contained in a k-deletion of w then $w(i) \geq 2$, but a_i can fail to be contained in a k-deletion of w for two reasons: either $w(i) = 1$ or $\|a_i\| > n - k$. To eliminate the latter possibility, let t denote the number of ≥2 entries in w. Because w has precisely k ones we have $\mathrm{ex}(w) \geq 2k + 1 - t$ and $\mathrm{ex}(w) \geq t$, which combine to show that $\mathrm{ex}(w) \geq k + 1$ for all values of t, so

$$n = m + \mathrm{ex}(w) \geq m + k + 1 = \|a_i\| + k.$$

Thus $\|a_i\| \leq n - k$, so a_i is contained in a k-deletion of w if and only if $w(i) \geq 2$, from which the ≥2 entries of w can be discerned, completing the proof. □

Example 6. Suppose the reconstruction algorithm is given the set of 3-deletions

$$\{322, 2212, 2221, 3112, 3121, 3211, 12121, 12211, 21121,\\ 21211, 22111, 31111, 111211, 121111, 211111\}.$$

of an unknown composition w of $n = 10$. Since the longest 3-deletions in this set are 3 letters longer than the shortest 3-deletion, w has at least $k = 3$ ones by Lemma 2. As the set also contains a 3-deletion without ones, the same lemma shows that w has precisely 3 ones, and thus the algorithm follows the proof of Lemma 5. The 3-deletion without ones — 322 — gives the $\geq$2 entries of w in their correct order. Now we form the a_i's to see where these $\geq$2 entries lie:

$a_1 = 211111$ is contained in a 3-deletion so $w(1) \geq 2$,
$a_2 = 121111$ is contained in a 3-deletion so $w(2) \geq 2$,
$a_3 = 112111$ is not contained in a 3-deletion so $w(3) = 1$,
$a_4 = 111211$ is contained in a 3-deletion so $w(4) \geq 2$,
$a_5 = 111121$ is not contained in a 3-deletion so $w(5) = 1$,
$a_6 = 111112$ is not contained in a 3-deletion so $w(6) = 1$.

Therefore we get $w = 321211$.

This leaves us to consider the case of compositions with many ones. In this case we also need the *second exceedance number*, defined by $\mathrm{ex}_2(w) = \sum(w(i) - 2)$ where the sum is over all entries $w(i) \geq 2$.

Lemma 7. *If w is a composition of $n \geq 3k + 1$ with more than k ones, then w can be reconstructed from its set of k-deletions.*

Proof. Given the set of k-deletions of such a composition w, our algorithm can apply the result of Lemma 2 to conclude that it has more than k ones. Therefore the k-deletions with the fewest ones contain all $\geq$2 entries of w in the order in which they occur in w; let $v = v(1) \cdots v(\ell)$ denote the composition formed by these entries, so

$$w = \underbrace{1 \cdots 1}_{z(1)} v(1) \underbrace{1 \cdots 1}_{z(2)} v(2) \cdots v(\ell - 1) \underbrace{1 \cdots 1}_{z(\ell)} v(\ell) \underbrace{1 \cdots 1}_{z(\ell+1)}$$

for some word $z \in \mathbb{N}^{\ell+1}$ (we take $\mathbb{N}$ to denote the nonnegative integers). Our goal is thus to determine z. We use similar techniques as in the proof of Lemma 3, although here we must perform two steps.

The first of these steps is to find the zeros in z. For $1 \leq i \leq \ell + 1$ let

$$a_i = \underbrace{2 \cdots 2}_{i-1} 1 \underbrace{2 \cdots 2}_{\ell+1-i}.$$

Since the twos in a_i can only embed into $\geq$2's in w, if a_i is contained in a k-deletion of w then its 1 must embed into an element between $v(i - 1)$ and $v(i)$, implying that $z(i) \geq 1$.

Conversely, if a_i is not contained in a k-deletion of w then either $\|a_i\| > n-k$ or $z(i) = 0$. Simple accounting shows that

$$n - k = (\text{\# of ones in } w) + 2\ell + \mathrm{ex}_2(w) - k,$$

so $\|a_i\| = 2\ell + 1 \le n - k$ because w has more than k ones, and thus

$$z(i) = 0 \iff a_i \text{ is not contained in a } k\text{-deletion of } w.$$

The second step is to use these zeros to divine the nonzero entries of z. Define the composition $b_i = b_i(1)\cdots b_i(\ell)$ by

$$b_i(j) = \begin{cases} 1 & \text{if} \quad j \le i-1 \text{ and } z(j) = 0 \text{ or} \\ & \qquad j \ge i \text{ and } z(j+1) = 0\text{, or} \\ 2 & \text{otherwise,} \end{cases}$$

and consider the possible embeddings of b_i in w. Suppose for the sake of example that $i \ge 4$. If $z(1) \ge 1$ then $b_i(1) = 2$ and thus can embed only into or to the right of $v(1)$. Otherwise if $z(1) = 0$ then $b_i(1) = 1$, but in this case $v(1)$ is the first entry of w so again $b_i(1)$ can embed only into or to the right of $v(1)$. If $z(2) \ge 1$ then $b_i(2) = 2$, and since $b_i(2)$ can only embed into a $\ge$2 entry in w to the right of $b_i(1)$, $b_i(2)$ can only embed into or to the right of $v(2)$. Otherwise if $z(2) = 0$ then $b_i(2) = 1$, but then $v(1)$ and $v(2)$ are adjacent in w so since $b_i(1)$ must embed into or to the right of $v(1)$ and $b_i(2)$ must embed to the right of $b_i(1)$ we see that $b_i(2)$ must embed into or to the right of $v(2)$. Continuing in this manner it is easy to see (or more formally, to prove inductively) that:

- For all $j \le i-1$, $b_i(j)$ must embed into or to the right of $v(j)$.
- For all $j \ge i$, $b_i(j)$ must embed into or to the left of $v(j)$.

These two facts combine to show that $b_i(i-1)$ and $b_i(i)$ can only embed between $v(i-1)$ and $v(i)$, inclusive. Now define the word $x \in \mathbb{N}^{\ell+1}$ by $x(i) = 0$ if $z(i) = 0$ and otherwise

$$x(i) = \max\{s : b_i(1)\cdots b_i(i-1)\underbrace{1\cdots 1}_{s}b_i(i)\cdots b_i(\ell) \text{ is contained in a } k\text{-deletion of } w\}.$$

The analogue to (1) now follows by the conditions on embeddings of b_i established above:

$$x(i) = \min\{z(i), n-k-\|b_i\|\}. \tag{4}$$

Suppose $z(i) \ge 1$. In this case $\|b_i\| = 2\ell - h$, where h denotes the number of 0 entries of z ("holes"). Letting $k+t$ denote the number of ones in w, we have

$$n = k + t + 2\ell + \mathrm{ex}_2(w),$$

allowing us to rewrite (4) as

$$x(i) = \min\{z(i), h + t + \mathrm{ex}_2(w)\}. \tag{5}$$

Paralleling the proof of Lemma 3, there are now two cases in which we are done:

- If $\|v\| + \|x\| = n$ then we must have $z = x$ and thus have successfully reconstructed w; by (5) this will occur if z contains no entries greater than $h + t + \mathrm{ex}_2(w)$.
- If at most one entry of x satisfies $x(i) = h + t + \mathrm{ex}_2(w)$ then by (5) we must have $z(j) = x(j)$ for all $j \neq i$, and then $z(j)$ can be calculated from the fact that $\|w\| = n$.

Suppose, for the sake of contradiction, that neither of these conditions holds, so z contains an entry $z(i) \geq h + t + \mathrm{ex}_2(w) + 1$ and another entry $z(j) \geq h + t + \mathrm{ex}_2(w)$. As each of the other $(\ell + 1 - h) - 2$ nonzero entries of z correspond to at least one one in w, we have

$$k + t = \text{\# of ones in } w \geq \ell + h + 2t + 2\,\mathrm{ex}_2(w).$$

From this it follows that

$$2k \geq t + 2\ell + \mathrm{ex}_2(w),$$

so

$$3k \geq k + t + 2\ell + \mathrm{ex}_2(w) = n,$$

and this contradiction completes the proof of both the lemma and Theorem 1. □

Example 8. Suppose the reconstruction algorithm is given the set of 3-deletions

$$\{1222, 2212, 11122, 11212, 11221, 12112, 12211, 111112, 111121, 111211, 112111, 1111111\}.$$

of an unknown composition w of $n = 10$. This set contains $1^{10-3} = 1111111$ and every 3-deletion in the set contains a 1, so Lemma 2 shows that w has more than $k = 3$ ones. Thus we follow the proof of Lemma 7. Each of the compositions with the fewest ones, e.g., 1222, give the $\geq$2 entries of w in their correct order, $v = 222$, so

$$w = \underbrace{1 \cdots 1}_{z(1)} 2 \underbrace{1 \cdots 1}_{z(2)} 2 \underbrace{1 \cdots 1}_{z(3)} 2 \underbrace{1 \cdots 1}_{z(4)}.$$

We then find the 0 entries of z:

$z(1) \neq 0$ because $a_1 = 1222$ is contained in a 3-deletion of w,
$z(2) = 0$ because $a_2 = 2122$ is not contained in a 3-deletion of w,
$z(3) \neq 0$ because $a_3 = 2212$ is contained in a 3-deletion of w,
$z(4) = 0$ because $a_4 = 2221$ is not contained in a 3-deletion of w.

Now we build the word $x \in \mathbb{N}^4$. We have that $x(2) = x(4) = 0$ because the corresponding entries of z are 0. To compute the other entries of x we construct $b_1 = 121$ and $b_3 = 211$ and then have

$x(1) = 3$ because 111 121 is contained in a 3-deletion of w but 1111 121 is not,
$x(3) = 1$ because 21 1 1 is contained in a 3-deletion of w but 21 11 1 is not.

Since $\|v\| + \|x\| = \|222\| + \|3010\| = 10$, we must have $z = x$ and thus $w = 1112212$.

The connection to permutations. The subject of permutation patterns (see Bóna's text [3] for a survey) is concerned with the following partial order on permutation: for permutations σ of length k and π of length n, let $\sigma \leq \pi$ if there are indices $i_1 < i_2 < \cdots < i_k$ such that the subsequence $\pi(i_1)\pi(i_2)\cdots\pi(i_k)$ has the same pairwise comparisons as $\sigma(1)\sigma(2)\cdots\sigma(k)$, and in such a case σ is said to be an $(n-k)$-deletion of π. For example, $13254 \leq 213654798$ (note that we write permutations in one-line, or list, notation) because of the subsequence 26598 $(= \pi(1)\pi(4)\pi(5)\pi(8)\pi(9))$.

Given two permutations σ and π of lengths m and n respectively, their *direct sum*, $\sigma \oplus \pi$, is the permutation of length $m+n$ whose first m entries form σ and whose last n entries are the copy of π obtained by adding m to each entry. For example, $213654 \oplus 132 = 213654798$. A permutation is said to be *layered* if it can be written as the direct sum of decreasing permutations. Thus 213654798 is layered because it can be written as $21 \oplus 1 \oplus 321 \oplus 1 \oplus 21$. There is a natural order-preserving bijection between layered permutations and compositions; for example, $213654798 = 21 \oplus 1 \oplus 321 \oplus 1 \oplus 21$ maps to the composition 21312 while $13254 = 1 \oplus 21 \oplus 21$ maps to 122, and $122 \leq 21312$ under the partial order on compositions.

Smith [11] was the first to study multiset reconstruction for permutations. Her work was followed by Raykova [9] who proved that for all k, all sufficiently long permutations are reconstructible from their multisets of k-deletions. This leaves open the question of whether all sufficiently long permutations are reconstructible from their *sets* of k-deletions. Our work answers Raykova's question of whether all sufficiently long layered permutations can be reconstructed from their sets of k-deletions.

Acknowledgement. I thank Robert Brignall and the referees for their helpful comments.

REFERENCES

[1] BERGERON, F., BOUSQUET-MÉLOU, M., AND DULUCQ, S. Standard paths in the composition poset. *Ann. Sci. Math. Québec 19*, 2 (1995), 139–151.

[2] BJÖRNER, A., AND SAGAN, B. Rationality of the Möbius function of a composition poset. *Theoret. Comput. Sci. 359*, 1-3 (2006), 282–298.

[3] BÓNA, M. *Combinatorics of permutations*. Discrete Mathematics and its Applications (Boca Raton). Chapman & Hall/CRC, Boca Raton, FL, 2004. With a foreword by Richard Stanley.

[4] CAMERON, P. J. Stories from the age of reconstruction. *Congr. Numer. 113* (1996), 31–41.

[5] HARARY, F. On the reconstruction of a graph from a collection of subgraphs. In *Theory of Graphs and its Applications (Proc. Sympos. Smolenice, 1963)*. Publ. House Czechoslovak Acad. Sci., Prague, 1964, pp. 47–52.

[6] LOTHAIRE, M. *Combinatorics on words*, vol. 17 of *Encyclopedia of Mathematics and its Applications*. Addison-Wesley Publishing Co., Reading, Mass., 1983.

[7] MANVEL, B. Some basic observations on Kelly's conjecture for graphs. *Discrete Math. 8* (1974), 181–185.

[8] PRETZEL, O., AND SIEMONS, J. On the reconstruction of partitions and applications. *Electron. J. Combin. 11 (2)* (2005), Note 5, 6 pp. (electronic).

[9] RAYKOVA, M. Permutation reconstruction from minors. *Electron. J. Combin. 13* (2006), Research paper 66, 14 pp. (electronic).

[10] SAGAN, B. E., AND VATTER, V. The Möbius function of a composition poset. *J. Algebraic Combin. 24*, 2 (2006), 117–136.

[11] SMITH, R. Permutation reconstruction. *Electron. J. Combin. 13* (2006), Note 11, 8 pp. (electronic).

[12] SNELLMAN, J. Saturated chains in composition posets. arXiv:math.CO/0505262.

[13] SNELLMAN, J. Standard paths in another composition poset. *Electron. J. Combin. 11*, 1 (2004), Research Paper 76, 8 pp. (electronic).