Generalisations of the binomial coefficient of words

Antoine Renard[1] Michel Rigo[1] Markus A. Whiteland[2]

[1]University of Liège [2]Loughborough University

q-deformed binomial coefficient of words

The ***binomial coefficient*** $\binom{u}{v}$ of two words $u, v \in A^*$ counts the number of occurrences of v as a scattered subword of u.

The ***q-deformation*** $\binom{\cdot}{\cdot}_q$ of the binomial coefficient of words is a polynomial in $\mathbb{N}[q]$ defined as follows: for all $u, v \in A^*$ and $a, b \in A$,

$$\binom{u}{\varepsilon}_q = 1, \quad \binom{\varepsilon}{v}_q = 0 \text{ if } v \neq \varepsilon,$$

$$\binom{ua}{vb}_q = \binom{u}{vb}_q \cdot q^{|vb|} + \delta_{a,b}\binom{u}{v}_q.$$

Examples of coefficients

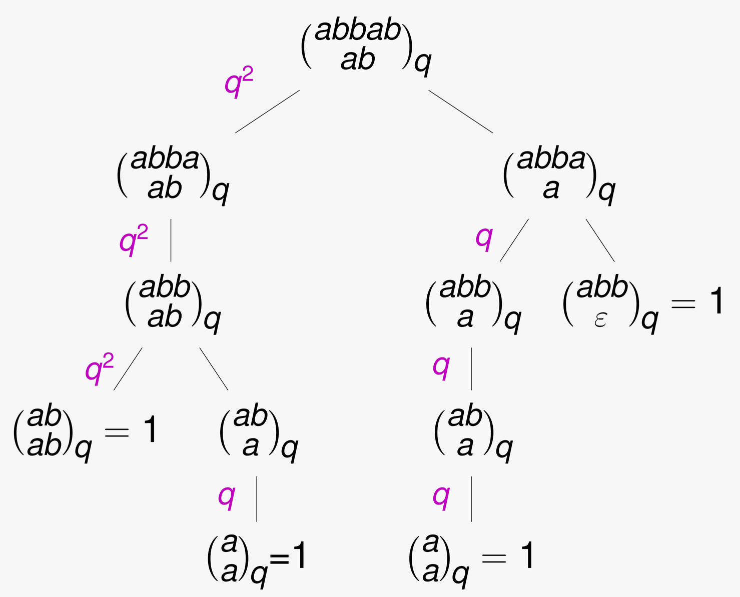
For the classical coefficient, we have

$$\binom{abbab}{ab} = 4 \quad \text{because} \quad \begin{matrix} abbab & abbab \\ abbab & abbab \end{matrix}.$$

The corresponding q-deformed coefficient is given by

$$\binom{abbab}{ab}_q = q^6 + q^5 + q^3 + 1$$

and can be computed using a tree:



Combinatorial interpretation

We get information about the positions of the different occurrences of v in u:

Theorem (R., Rigo, Whiteland, 2025). Let $u \in A^*$, $k \geq 0$, and $a_1, \ldots, a_k \in A$. Then

$$\binom{u}{a_1 \cdots a_k}_q = \sum_{\substack{u_0, u_1, \ldots, u_k \in A^* \\ u = u_0 a_1 \cdots u_{k-1} a_k u_k}} q^{\sum_{i=1}^{k} i|u_i|}.$$

In other words, each occurrence of $v = a_1 \cdots a_k$ in u contributes to $\binom{u}{v}_q$ with a term q^α, where α is the sum over all letters of v of the number of letters to the right of them and that are not part of this particular occurrence of v.

The same example, another point of view

Considering the previous example, we have

$$\begin{array}{ccccccccc} ab\overline{\underline{bab}} & & a\underline{b}b\overline{\underline{ab}} & & a\underline{bba}b & & abbab & & \\ \downarrow & & \downarrow & & \downarrow & & \downarrow & & \\ q^{3+3} & + & q^{2+3} & + & q^{0+3} & + & q^{0+0} & = & \binom{abbab}{ab}_q \end{array}$$

Characterising p-group languages

A language $L \subset A^*$ is ***recognisable*** if and only if there exist

a finite monoid M, a subset $S \subset M$ and a monoid morphism $\varphi : A^* \to M$,

such that $L = \varphi^{-1}(S)$. A language recognised by a p-group is a ***p-group language***.

Theorem (Eilenberg, 1976). Let p be a prime. A language is a p-group language if and only if it is a Boolean combination of languages of the form

$$L_{v,r,p} := \left\{u \in A^* \mid \binom{u}{v} \equiv r \pmod{p}\right\}.$$

We generalise this result using our q-deformed coefficients.

Theorem (R., Rigo, Whiteland, 2025). Let p be a prime. A language is a p-group language if and only if it is a Boolean combination of languages of the form

$$L_{v,\mathfrak{R},p} = \{u \in A^* : \binom{u}{v}_q \equiv \mathfrak{R} \pmod{q^p - 1}\}.$$

q-Parikh matrices

Let $z = z_1 \cdots z_n$ be a word and A be the alphabet of z, *i.e.* the set of letters occurring in z. For $a \in A$ and $\ell \geq 0$, we let $\mathcal{M}_{a,\ell}$ denote the upper triangular matrix defined by

$$[\mathcal{M}_{a,\ell}]_{i,j} = \begin{cases} 1 & j = i \\ \delta_{a,z_i} q^\ell & j = i+1 \\ 0 & \text{otherwise} \end{cases}.$$

We define the ***q-Parikh matrix mapping*** as follows:

$$\mathcal{P}_z : A^* \to (\mathbb{N}[q])^{(n+1)\times(n+1)} : u_k u_{k-1} \cdots u_1 u_0 \mapsto \mathcal{M}_{u_k,k} \cdots \mathcal{M}_{u_0,0}.$$

Link between q-deformed coefficients and q-Parikh matrices

Unlike the classical Parikh matrix mapping, $\mathcal{P}_z$ is not a monoid morphism. Nonetheless, the following result establishes a connection with our q-binomial coefficient of words.

Theorem (R., Rigo, Whiteland, 2025). Let z be a word of length $n \geq 1$ whose alphabet is A. Let $u \in A^*$. The corresponding $(n+1) \times (n+1)$ q-Parikh matrix is such that

- ⋆ $[\mathcal{P}_z(u)]_{i,j} = 0$, for all $1 \leq j < i \leq n+1$,
- ⋆ $[\mathcal{P}_z(u)]_{i,i} = 1$, for all $1 \leq i \leq n+1$.
- ⋆ Let $r \in \{1, \ldots, n\}$. For all $1 \leq i \leq n - r + 1$, $[\mathcal{P}_z(u)]_{i,i+r} = q^{\mathrm{s}(r-1)}\binom{u}{z_i z_{i+1} \cdots z_{i+r-1}}_q$, where $\mathrm{s}(\ell) = \sum_{k=1}^{\ell} k$.

A q-Parikh matrix

Let $z = aba$ and $u = abbaba$, we have

$$\mathcal{M}_{a,\ell} = \begin{pmatrix} 1 & q^\ell & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & q^\ell \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad \mathcal{M}_{b,\ell} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & q^\ell & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

so that

$$\mathcal{P}_z(u) = \begin{pmatrix} 1 & \binom{u}{a}_q & q\binom{u}{ab}_q & q^3\binom{u}{aba}_q \\ 0 & 1 & \binom{u}{b}_q & q\binom{u}{ba}_q \\ 0 & 0 & 1 & \binom{u}{a}_q \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

q-coefficients of infinite periodic words

Theorem (R., Rigo, Whiteland, 2025). The q-binomial $\binom{u^n}{z}_q$ can be expressed as

$$\frac{1}{q^{\binom{|z|}{2}}} \sum_{k=1}^{m} R_k(q) \frac{1 - q^{c_k n|u|}}{1 - q^{c_k|u|}},$$

where m and c_k are positive integers, and R_k are rational functions whose denominators only have factors of the form $(1 - q^{t|u|})$ for some integer t. Moreover, these quantities c_k and R_k can be effectively computed. In particular, the sequence $(\binom{u^n}{z}_q)_{n\geq 0}$ converges in $\mathbb{N}[[q]]$ to the formal power series $\mathfrak{s}_{\mathbf{u},z}(q)$ expressed by the rational function

$$\frac{1}{q^{\mathrm{s}(|z|-1)}} \sum_{k=1}^{m} R_k(q) \frac{1}{1 - q^{c_k|u|}}.$$

An example of such a series

The sequence $\left(\binom{(abba)^n}{ab}_q\right)_{n\geq 0}$ converges to the series

$$q^3 + 2q^4 + q^5 + q^7 + 2q^8 + q^9 + 2q^{11} + 4q^{12} + 2q^{13} + 2q^{15} + 4q^{16} + 2q^{17} + 3q^{19} + 6q^{20} + \cdots,$$

which corresponds to the rational function

$$R(q) = \frac{q^3}{(q-1)^2 (q^2+1)^2 (q^4+1)}.$$

Towards a binomial coefficient over the free group

The binomial coefficient of words can also be defined through the ***Magnus transform***:

$$\mu : A^* \to U(\mathbb{Z}\langle\langle A \rangle\rangle), \quad \text{such that} \quad \mu(a) = 1 + a, \ \forall a \in A.$$

We then have

$$\mu(u) = \sum_{v \in A^*} \binom{u}{v} v.$$

In [4], the Magnus transform is extended to a group morphism:

$$\mu(a^{-1}) = (1+a)^{-1} = \sum_{n\geq 0} (-1)^n a^n.$$

We can thus define $\binom{u}{v}$ for $u \in F(A)$, $v \in A^*$. What can we say about it?

References

[1] Samuel Eilenberg.
Automata, languages, and machines. Vol. B.
Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1976.
Pure and Applied Mathematics, Vol. 59.

[2] Antoine Renard, Michel Rigo, and Markus A. Whiteland.
Introducing q-deformed binomial coefficients of words.
J. Algebraic Combin., 61(2):25, 2025.

[3] Antoine Renard, Michel Rigo, and Markus A. Whiteland.
q-Parikh matrices and q-deformed binomial coefficients of words.
Discrete Math., 348(5):Paper No. 114381, 21, 2025.

[4] Christophe Reutenauer.
Free Lie algebras, volume 7 of Lond. Math. Soc. Monogr., New Ser.
The Clarendon Press, Oxford University Press, New York, 1993.
Oxford Science Publications.