Merging tensor views

Niels uit de Bos

March 5, 2025

This document contains a formalization of the concept of merging tensor views. We prove a simple criterion for when two tensor views can be merged into a single view without changing the underlying memory layout: the main result (theorem 10) gives necessary and sufficient conditions for when two views can be merged, based on checking additivity conditions of the composed index functions. A worked out example in section 4 illustrates why these conditions are necessary and cannot be simplified further.

The code and formal proofs in the Lean theorem prover are available at `https://github.com/nielius/tensorlayouts`.

1 Views

Definition 1. A *view* is a tuple (s, σ) of a shape $s = (s_1, \ldots, s_k)$, $s_i \in \mathbb{N}_{\geq 2}$, and a tuple of strides $\sigma = (\sigma_1, \ldots, \sigma_k)$, $\sigma_i \in \mathbb{N}$.

Note that we only consider sizes ≥ 2: a size of 1 or 0 is empty and can be ignored.

Definition 2. The *index set* of a shape $s = (s_1, \ldots, s_k) \in \mathbb{N}^k$ is the set

$$I_s = \{(i_1, \ldots, i_k) \in \mathbb{N}^k : \forall 1 \leq \ell \leq k, 0 \leq i_\ell < s_\ell\}.$$

Definition 3. The *index function* of a view $v = (s, \sigma)$ is the function

$$v \colon I_s \to \mathbb{N}, \quad (i_1, \ldots, i_k) \mapsto \sum_\ell i_\ell \sigma_\ell.$$

Note that we abuse notation: we use the same symbol v for the view and the index function.

The additivity of the inner product in the definition above gives us the following lemma.

Lemma 4. *Let $v = (s, \sigma)$ be a view. Then v is additive in I_s, i.e.,*

$$v(i + i') = v(i) + v(i') \quad \textit{for all } i, i' \in I_s \textit{ with } i + i' \in I_s.$$

1

Definition 5. Let $s = (s_1, \ldots, s_k)$ be a shape. The *contiguous view* for s is the view $v^s = (s, \sigma^s)$, where

$$\sigma_j^s = \prod_{\ell=j+1}^{k} s_\ell.$$

This is the standard row-major view that a tensor of shape s has.

The interpretation of these views is as follows: with respect to some given contiguous and linear array of data d, a view $v = (s, \sigma)$ represents a tensor of shape s, with the coordinate $[i_1, \ldots, i_k]$ corresponding to the data point $d\left[v(i_1, \ldots, i_k)\right] = d\left[\sum_j i_j \sigma_j\right]$.

2 Merging Views

Definition 6. Let $v_1 = (s_1, \sigma_1), v_2 = (s_2, \sigma_2)$ be views. We define the *functional composition* of v_1 and v_2 as the function

$$f_{v_1, v_2} := v_1 \circ (v^{s_1})^{-1} \circ v_2 \colon I_{s_2} \to \mathbb{N}.$$

Remark 7. For a view $v = (s, \sigma)$, the function $(v^s)^{-1}$ is also known as the *unraveling* function: it takes a natural number i and interprets it as an index of a tensor of shape s, where the tensor layout is assumed to be row-major, i.e., the standard contiguous layout. We need to insert this unraveling function in the definition of the functional composition, because we want to think of v_2 as a view into a tensor of shape s_1, but v_2 as defined so far is a function $I_{s_2} \to \mathbb{N}$.

Definition 8. Let $v_1 = (s_1, \sigma_1), v_2 = (s_2, \sigma_2), v = (s, \sigma)$ be views. We say that v is the *merged view* of v_1 and v_2 if

$$v = f_{v_1, v_2}$$

as functions $I_{s_2} \to \mathbb{N}$. (Note that this implies $s = s_2$.)

Definition 9. Let $v_1 = (s_1, \sigma_1), v_2 = (s_2, \sigma_2), v = (s, \sigma)$ be views. We say that (v_1, v_2) is *mergeable* or *can be merged* if there exists a view v such that v is the merged view of v_1 and v_2.

The main result of this paper is necessary and sufficient conditions for two views to be mergeable.

3 Main result

Let $v = (s = (s_1, \ldots, s_k), \sigma = (\sigma_1, \ldots, \sigma_k))$ be a view and let $1 \le j \le k$ be an integer. We will denote by $e_j = (0, \ldots, 0, 1, 0, \ldots, 0) \in \mathbb{N}^k$ the vector with a 1 in the j-th position and 0 elsewhere. We can add e_j to an index $i \in I_s$ if $i_j + 1 < s_j$; this gives us an index $i + e_j \in I_s$.

We recall (definition 6) that for views v, v', we denote the functional composition of those views by $f_{v,v'}$; (v, v') is defined to be mergeable (definition 9) if there exists a view $\tilde{v}$ such that the index function of $\tilde{v}$ (definition 3; abusing notation to also denote this index function by $\tilde{v}$) is equal to $f_{v,v'}$.

Theorem 10. *Let $v = (s = (s_1, \ldots, s_k), \sigma = (\sigma_1, \ldots, \sigma_k))$ and $v' = (s' = (s_1, \ldots, s_\ell), \sigma' = (\sigma'_1, \ldots, \sigma'_\ell))$ be views. Then (v, v') is mergeable if and only if for all $i \in I_{s'}$ and all $1 \leq j \leq \ell'$ such that $i_j + 1 < s'_j$, we have*

$$f_{v,v'}(i + e_j) = f_{v,v'}(i) + f_{v,v'}(e_j).$$

If (v, v') is mergeable, then the merged view $\tilde{v}$ is unique and equal to

$$\tilde{v} = (s', (\tilde{\sigma}_1, \ldots, \tilde{\sigma}_\ell)), \quad where \quad \tilde{\sigma}_j = f_{v,v'}(e_j).$$

A Lean formalization of this theorem is available at `https://github.com/nielius/tensorlayouts`.

Remark 11. This theorem on its own is not very surprising. Perhaps it is more interesting to note that this can not be simplified further, and that you actually need to check the conditions for all $i \in I_{s'}$ and all $1 \leq j \leq \ell'$ such that $i_j + 1 < s'_j$. This is made clearer by the example in section 4.

Remark 12. The criterion given in theorem 10 is easy to check computationally, although it involves checking a large number of conditions.

Proof of theorem 10. ($\Rightarrow$) Assume that (v, v') is mergeable, i.e., there exists a view $\tilde{v} = (s', \tilde{\sigma})$ such that $\tilde{v} = f_{v,v'}$. Lemma 4 shows that $\tilde{v}$ is additive in $I_{s'}$, so in particular, we have the given condition.

($\Leftarrow$) Define $\tilde{\sigma} = (\tilde{\sigma}_1, \ldots, \tilde{\sigma}_\ell)$ by $\tilde{\sigma}_j = f_{v,v'}(e_j)$. We can now show $\tilde{v} = f_{v,v'}$ by induction:

- $\tilde{v}(0) = 0$ and $f_{v,v'}(0) = 0$ are clear.

- Assume that $\tilde{v}(i) = f_{v,v'}(i)$ for all $i \in I_{s'}$ with $\sum_j i_j \leq K$ for some given K. (The base case is $K = 0$, which we remarked on above.) Any index $i' \in I_{s'}$ with $\sum_j i'_j = K + 1$ can be written as $i' = i + e_j$ for some $i \in I_{s'}$ with $\sum_{j'} i_{j'} = K$ and $1 \leq j \leq \ell'$. Then we have

$$
\begin{aligned}
f_{v,v'}(i') = f_{v,v'}(i + e_j) &= f_{v,v'}(i) + f_{v,v'}(e_j) && \text{by assumption} \\
&= \tilde{v}(i) + \tilde{v}(e_j) && \text{by the induction hypothesis} \\
&= \tilde{v}(i') && \text{by lemma 4}
\end{aligned}
$$

This shows that $\tilde{v} = f_{v,v'}$, completing the proof. $\qquad\square$

4 Example: Checking mergeability through overflows

Let

$$v = ((10, 3, 3), (\sigma_1, \sigma_2, \sigma_3)) \qquad \text{(where } \sigma_1, \sigma_2, \sigma_3 \in \mathbb{N})$$
$$v' = ((s), (4)) \qquad \text{(where } s \in \mathbb{N})$$

be two views. (Note that the shape of v' is (s), i.e., it is a one-dimensional shape.) In this example, we will check our main criterion (theorem 10) to see for which values of s and $\sigma_1, \sigma_2, \sigma_3$ the views (v, v') are mergeable. This is a relatively simple example, because v' is one-dimensional, but it shows the main problems you encounter.

We will see that (v, v') is mergeable if $s \leq 4$ and $\sigma_1 = 2\sigma_2 + 3\sigma_3$ hold. However, if $s > 4$, then for (v, v') to be mergeable, we additionally require that $\sigma_1 = 3\sigma_2$ holds. This shows that this problem is not simply a linear algebra problem, and the only way to determine the necessary conditions is to iterate over indices and analyze the overflow structure. The reason for this, as we will see, is that for $s > 4$, we have an additional overflow condition that needs to be checked.

This is the approach we are going to take. By definition, (v, v') can be merged if the functional composition

$$f_{v,v'} = v \circ (v^{(10,3,3)})^{-1} \circ v'$$

is equal to some view $\tilde{v} = (s, \tilde{\sigma})$. The main criterion (theorem 10) says that for (v, v') to be mergeable, there has to be a $\tilde{\sigma}$ such that

$$f_{v,v'}(i + 1) = f_{v,v'}(i) + \tilde{\sigma}, \qquad \text{or equivalently} \quad f_{v,v'}(i + 1) - f_{v,v'}(i) = \tilde{\sigma}. \quad (1)$$

for all $0 \leq i < s - 1$. To check this, we will simply evaluate $f_{v,v'}(i)$ for all $i \in I_{(s)} = \{0, \ldots, s - 1\}$, and note down the equation $f_{v,v'}(i + 1) - f_{v,v'}(i) = \tilde{\sigma}$ that we get.

4.1 Case $s = 4$

Proceeding as described above, we get the following table. Let's write out the steps for $i = 1$ as an example. We have $v'(1) = 4 \cdot 1 = 4$ (because the stride of v' is 4). The unraveling of 4 in the shape $(10, 3, 3)$ is $(0, 1, 1)$ (the easiest way to do this is to realize that the unravel function is like arithmetic in a mixed radix system), so $((v^{(10,3,3)})^{-1} \circ v')(1) = (0, 1, 1)$. Lastly, we have $f_{v,v'}(1) = v(0, 1, 1) = 0 \cdot \sigma_1 + 1 \cdot \sigma_2 + 1 \cdot \sigma_3 = \sigma_2 + \sigma_3$ Since $f_{v,v'}(0) = 0$, we get the equation

$$\sigma_2 + \sigma_3 = \tilde{\sigma}.$$

This is what we show in the row $i = 1$ of the table. The calculations for the other rows are similar.

i	i_1	i_2	i_3	Equation
0	0	0	0	
1	0	1	1	$\tilde{\sigma} = \sigma_2 + \sigma_3$
2	0	2	2	$\tilde{\sigma} = \sigma_2 + \sigma_3$
3	1	1	0	$\sigma_1 + \sigma_2 = 2\sigma_2 + 2\sigma_3 + \tilde{\sigma}$

Table 1: Overflow equations for $s = 4$. For every value of i, we calculate $(i_1, i_2, i_3) = ((v^{(10,3,3)})^{-1} \circ v')(i)$; these are the coordinates in the tensor v that correspond to the index i in the tensor v'. The last column, labelled with Equation, shows the equation we get from the main criterion (eq. (1)).

The most interesting row is the last one: it is where index i_2 overflows into index i_1, and index i_3 overflows into index i_2. As a result of this, we get a different equation than for the other rows. (This is a general phenomenon: every overflow gives you a new equation, and the precise equation depends only on which indices overflowed.) From the last row, we obtain the equation $\sigma_1 = 2\sigma_2 + 3\sigma_3$. Solving these equations shows that (v, v') is mergeable if and only if $\sigma_1 = 2\sigma_2 + 3\sigma_3$. The merged view is then equal to $((4), (\sigma_2 + \sigma_3))$.

4.2 Case $s = 6$

Now consider $s = 6$. We get the following equations:

i	i_1	i_2	i_3	Equation
0	0	0	0	
1	0	1	1	$\tilde{\sigma} = \sigma_2 + \sigma_3$
2	0	2	2	$\tilde{\sigma} = \sigma_2 + \sigma_3$
3	1	1	0	$\sigma_1 + \sigma_2 = 2\sigma_2 + 2\sigma_3 + \tilde{\sigma}$
4	1	2	1	$\tilde{\sigma} = \sigma_2 + \sigma_3$
5	2	0	2	$2\sigma_1 + 2\sigma_3 = \sigma_1 + 2\sigma_2 + \sigma_3 + \tilde{\sigma}$

Table 2: Overflow equations for $s = 6$

The additional equation $2\sigma_1 + 2\sigma_3 = \sigma_1 + 2\sigma_2 + \sigma_3 + \sigma$ simplifies to $\sigma_1 = 3\sigma_2$. Thus, in the case $s = 6$, (v, v') is be mergeable if and only if $\sigma_1 = 3\sigma_2$ and $\sigma_1 = 2\sigma_2 + 3\sigma_3$ hold. We can rewrite these two equations to $\sigma_1 = 3\sigma_2$ and $\sigma_2 = 3\sigma_3$.

4.3 Observations

The most important conclusion from this example is that it is not enough to have some linear algebraic condition on the strides. The mergeability depends on the precise overflow structure, which in turn depends on the shape of the view. Bigger shapes are likely to encounter more overflows: as an extreme case, (v, v') is always mergeable if v' has shape (2).

- The generic case involves only a single overflow, while a double overflow imposes additional constraints.

- If all single overflows satisfy the required conditions, then more complex overflows will automatically hold.

- Overflows correspond to linear equations that must be checked to determine mergeability.

- This calculation is akin to arithmetic overflow in numeral systems, but generalized to mixed-radix numbering where each index position has a different base given by the shape.

This demonstrates that the only way to determine the necessary conditions is to iterate over indices and analyze the overflow structure.

If anyone knows of an efficient method to calculate which overflows will occur, please send me a message.

5 Other results

NOTE: THIS SECTION IS A DRAFT

- The unraveling function (as defined in remark 7) $(v^s)^{-1}$ for a shape s is the function that sends a natural number $i \in \mathbb{N}$ to its representation in a mixed radix numerical system. For example, if $s = (2, 3, 4)$, then $(v^s)^{-1}(5) = (0, 1, 1)$.

- In the main criterion (theorem 10), instead of checking all possible increments, you could check only the increments that lead to overflows (in the mixed radix arithmetic). Every overflow gives a linear equation in the strides of the first view in the merge. This equation only depends on the index components that are involved in the overflow. Potentially, you could use this to check fewer conditions.