COMP/LING 445/645
Problem Set 4

This problem set consists only of questions involving mathematics or English or or a combination of the two. Please put your answers in an **Answer** section like in the example below. You can find more information about LaTeX here https://www.latex-project.org/.

Once you have answered the question, please compile your copy of this LaTeX document into a PDF and submit (i) the compiled PDF (ii) the raw LaTeX file and (iii) your `pittard-victoria.clj` file via email to *both* timothy.odonnell@mcgill.ca and savanna.willerton@mail.mcgill.ca. The problem is set is due before 16:05 on Monday, November 25, 2019.

Problem 0: This is an example question using some fake math like this $L = \sum_0^\infty \mathcal{G}\delta_x$.

Answer 0: Put your answer right under the question like this $L = \sum_0^\infty \mathcal{G}\delta_x$.

Problem 1: In class we gave the following equation for the bigram probability of a sequence of words $W^{(1)}, \ldots, W^{(k)}$:

$$\Pr(W^{(1)}, \ldots, W^{(k)}) = \prod_i^k \Pr(W^{(i)}|W^{(i-1)} = w^{(i-1)}) \tag{1}$$

Using this formula, give an expression for the bigram probability of the sentence *abab*, where each character is treated as a word. Try to simplify the formula as much as possible.

Answer 1: In the sentence *abab*, the first and third words are a, while the second and fourth words are b. Also note that the probability that the first symbol in any sentence $(W^{(0)})$ is $\rtimes$ is 1. As such, the following math holds true:

$$Pr(W^{(1)}, \ldots, W^{(k)}) = \prod_i^k \Pr(W^{(i)}|W^{(i-1)} = w^{(i-1)})$$

$$Pr(W^{(1)}, \ldots, W^{(4)}) = \prod_i^4 \Pr(W^{(i)}|W^{(i-1)} = w^{(i-1)})$$

$$= Pr(W^{(1)}|W^{(0)} = \rtimes) * Pr(W^{(2)}|W^{(1)} = w^{(1)}) * Pr(W^{(3)}|W^{(2)} = w^{(2)}) * Pr(W^{(4)}|W^{(3)} = w^{(3)})$$

$$= Pr(W^{(1)}) * Pr(W^{(2)}|W^{(1)} = w^{(1)}) * Pr(W^{(3)}|W^{(2)} = w^{(2)}) * Pr(W^{(4)}|W^{(3)} = w^{(3)})$$

$$Pr(abab) = Pr(W^{(1)} = a, W^{(2)} = b, W^{(3)} = a, W^{(4)} = b)$$

$$= Pr(W^{(1)} = a)Pr(W^{(2)} = b|W^{(1)} = a)Pr(W^{(3)} = a|W^{(2)} = b)Pr(W^{(4)} = b|W^{(3)} = a)$$

Because the probability of any word $W^{(i)}$ depends solely on the probability of $W^{(i-1)}$, the probability $Pr(W^{(2)} = b|W^{(1)} = a)$ and the probability $Pr(W^{(4)} = b|W^{(3)} = a)$ are both simplified to the probability $Pr(W^{(i)} = b|W^{(i-1)} = a)$. Thus:

$$Pr(abab) = Pr(W^{(1)} = a) * Pr(W^{(i)} = b|W^{(i-1)} = a)^2 * Pr(W^{(i)} = a|W^{(i-1)} = b)$$

Problem 2: Suppose that there are two possible symbols/words in our language, a and b. There are three conditional distributions in the bigram model for this language, $\Pr(W^{(i)}|W^{(i-1)} = a)$, $\Pr(W^{(i)}|W^{(i-1)} = b)$, and $\Pr(W^{(i)}|W^{(i-1)} = \rtimes)$. These conditional distributions are associated with the parameter vectors $\vec{\theta}_a$, $\vec{\theta}_b$, and $\vec{\theta}_\rtimes$, respectively (these parameter vectors were implicit in the previous problem). For the current problem, we will assume that these parameters are fixed.

Suppose that we are given a sentence $W^{(1)}, \ldots, W^{(k)}$. We will use the notation $n_{x \to y}$ to denote the number of times that the symbol y occurs immediately following the symbol x in the sentence. For example, $n_{a \to a}$ counts the number of times that symbol a occurs immediately following the symbol a. Using Equation 1, give an expression for the probability of a sentence in our language:

$$\Pr(W^{(1)}, \ldots, W^{(k)}|\vec{\theta}_a, \vec{\theta}_b, \vec{\theta}_\rtimes)$$

The expression should make use of the $n_{x \to y}$ notation defined above.

Hint: the expression should be analogous to the formula that we found for the likelihood of a corpus under a bag of words model.

Answer 2: For this answer, note that the notation $\vec{\theta}_x[y]$ refers to the element in $\vec{\theta}_x$ (as defined in the problem) storing the probability that y occurs at i given x at i-1. Also note that as $\rtimes$ is the start sentence operator, it only occurs at the start of a sentence. Thus, it will never follow a or b. As such, we are only concerned with the probability of a after a, a after b, a after start, b after a, b after b, and b after start. Furthermore, note that these events occur $n_{x \to y}$ times, as specified in the problem. According to Equation 1 and rules of probability, this means that to find the probability that the event occurs n times, each probability $\vec{\theta}_x[y]$ is multiplied by itself $n_{x \to y}$ times; in other words, $\vec{\theta}_x[y]$ can be raised to the power of $n_{x \to y}$.

$$\Pr(W^{(1)}, \ldots, W^{(k)}|\vec{\theta}_a, \vec{\theta}_b, \vec{\theta}_\rtimes) = \Pr(W^{(1)}, \ldots, W^{(k)}) = \prod_{i}^{k} \Pr(W^{(i)}|W^{(i-1)} = w^{(i-1)}, \vec{\theta}_a, \vec{\theta}_b, \vec{\theta}_\rtimes)$$

given that for all i-1=a: $Pr(W^{(i)}) \sim \vec{\theta}_a$

for all i-1=b: $Pr(W^{(i)}) \sim \vec{\theta}_b$

for all i-1= $\rtimes$: $Pr(W^{(i)}) \sim \vec{\theta}_\rtimes$

$$\Pr(W^{(1)}, \ldots, W^{(k)}|\vec{\theta}_a, \vec{\theta}_b, \vec{\theta}_\rtimes) = \prod_{i \in \{a,b,\rtimes\}} \prod_{j \in \{a,b\}} \vec{\theta}_i[j]^{n_{i \to j}}$$

Problem 3: Assume the parameter vectors in our bigram model have the following values:

$$\vec{\theta}_a = (0.7, 0.3)$$
$$\vec{\theta}_b = (0.2, 0.8)$$
$$\vec{\theta}_\rtimes = (0.5, 0.5)$$

The first vector indicates that if the current symbol a, there is probability 0.7 of transitioning to the symbol a, and probability 0.3 of transitioning to the symbol b. Using your answer to the previous problem and these parameter values, calculate the probability of the string $aabab$.

Answer 3: Note that in the string *aabab*, *a* follows ⋈ once, *a* follows *a* once, *b* follows *a* twice, and *a* follows *b* once. Thus, $n_{⋈\to a} = 1$, $n_{a\to a} = 1$, $n_{a\to b} = 2$, $n_{b\to a} = 1$, and all other values of $n_{x\to y}$ equal 0.

$$\Pr(W^{(1)}, \ldots, W^{(k)} | \vec{\theta}_a, \vec{\theta}_b, \vec{\theta}_⋈) = \prod_{i\in\{a,b,⋈\}} \prod_{j\in\{a,b\}} \vec{\theta}_i[j]^{n_{i\to j}}$$

$$\Pr(aabab) = \Pr(W^{(1)} = a, W^{(2)} = a, W^{(3)} = b, W^{(4)} = a, W^{(5)} = b | \vec{\theta}_a, \vec{\theta}_b, \vec{\theta}_⋈)$$

$$= \vec{\theta}_⋈[a]^{n_{⋈\to a}} * \vec{\theta}_⋈[b]^{n_{⋈\to b}} * \vec{\theta}_a[a]^{n_{a\to a}} * \vec{\theta}_a[b]^{n_{a\to b}} * \vec{\theta}_b[a]^{n_{b\to a}} * \vec{\theta}_b[b]^{n_{b\to b}}$$

$$= \vec{\theta}_⋈[a]^1 * \vec{\theta}_⋈[b]^0 * \vec{\theta}_a[a]^1 * \vec{\theta}_a[b]^2 * \vec{\theta}_b[a]^1 * \vec{\theta}_b[b]^0$$

$$= \vec{\theta}_⋈[a] * \vec{\theta}_a[a] * \vec{\theta}_a[b]^2 * \vec{\theta}_b[a]$$

$$= .5 * .7 * .3^2 * .2$$

$$= .0063$$

Problem 4: In Problem 2, you found an expression for the bigram probability of a sentence in our language, which contains the symbols *a* and *b*. In that problem, we assumed that there were fixed parameter vectors $\vec{\theta}$ associated with each conditional distribution. In this problem, we will consider the setting in which we have uncertainty about the value of these parameters.

In class, we used the Dirichlet distribution to define a prior distribution over parameter vectors:

$$\vec{\theta}_{\mathbf{c}} \mid \vec{\alpha} \sim \qquad\qquad\qquad \text{Dirichlet}(\vec{\alpha}) \qquad\qquad (2)$$

$$w^{(i)} \mid w^{(i-1)} \sim \qquad\qquad\qquad \text{categorical}(\vec{\theta}_{w^{(i-1)}}) \qquad\qquad (3)$$

$$w^{(1)} \sim \qquad\qquad\qquad \text{categorical}(\vec{\theta}_⋈) \qquad\qquad (4)$$

Suppose that we have a sentence $S = W^{(1)}, \ldots, W^{(k)}$. Given an expression for the joint probability $\Pr(S, \vec{\theta}_a, \vec{\theta}_b, \vec{\theta}_⋈ | \vec{\alpha})$ using the definitions of Dirichlet distributions and likelihoods we defined in class.

Answer 4:

$$Pr(S, \vec{\theta}_a, \vec{\theta}_b, \vec{\theta}_⋈ | \vec{\alpha}) = Pr(S | \vec{\theta}_a, \vec{\theta}_b, \vec{\theta}_⋈) * Pr(\vec{\theta}_a | \vec{\alpha}) * Pr(\vec{\theta}_b | \vec{\alpha}) * Pr(\vec{\theta}_⋈ | \vec{\alpha})$$

$$= \left(\prod_{i\in\{a,b,⋈\}} \prod_{j\in\{a,b\}} \vec{\theta}_i[j]^{n_{i\to j}} \right) * \left(\prod_{i\in\{a,b,⋈\}} \vec{\theta}_i^{\vec{\alpha}-1} \right) * \frac{1}{B(\alpha)}$$

$$= \frac{\Gamma(\sum_{i\in\{a,b\}} \alpha_i)}{\prod_{i\in\{a,b\}} \Gamma(\alpha_i)} * \left(\prod_{i\in\{a,b,⋈\}} \prod_{j\in\{a,b\}} \vec{\theta}_i[j]^{n_{i\to j}} \right) * \left(\prod_{i\in\{a,b,⋈\}} \vec{\theta}_i^{\alpha_j - 1} \right)$$

$$= \frac{\Gamma(\sum_{i\in\{a,b\}} \alpha_i)}{\prod_{i\in\{a,b\}} \Gamma(\alpha_i)} * \left(\prod_{i\in\{a,b,⋈\}} \prod_{j\in\{a,b\}} \vec{\theta}_i[j]^{n_{i\to j} + \alpha_j - 1} \right)$$

Problem 5: In the previous problem, you found a formula for the joint probability of a sentence and a set of bigram model parameters. Using this, give a formula for the marginal probability of the sentence $\Pr(S | \vec{\alpha})$.

Hint: The formula should be analogous to the formula derived in class for marginal probability of a corpus under a bag of words model. Whereas before there was only a single parameter vector $\vec{\theta}$, now there are three parameters that need to be marginalized away. Otherwise the calculation will be similar.

Answer 5: To go from a joint to marginal distribution, the variables in the joint but not the marginal distribution have to be marginalized out. This is done by integrating over all values of the marginalized variables (in this case, the three thetas). Thus:

$$Pr(S|\vec{\alpha}) = \int_{\Theta_a} \int_{\Theta_b} \int_{\Theta_{\rtimes}} \frac{\Gamma(\sum_{i \in \{a,b\}} \alpha_i)}{\prod_{i \in \{a,b\}} \Gamma(\alpha_i)} * \left(\prod_{i \in \{a,b,\rtimes\}} \prod_{j \in \{a,b\}} \vec{\theta}_i[j]^{n_{i \to j} + \alpha_j - 1} \right)$$

$$= \frac{\Gamma(\sum_{i \in \{a,b\}} \alpha_i)}{\prod_{i \in \{a,b\}} \Gamma(\alpha_i)} \int_{\Theta_a} \int_{\Theta_b} \int_{\Theta_{\rtimes}} \left(\prod_{i \in \{a,b,\rtimes\}} \prod_{j \in \{a,b\}} \vec{\theta}_i[j]^{n_{i \to j} + \alpha_j - 1} \right)$$

$$= \frac{\Gamma(\sum_{i \in \{a,b\}} \alpha_i)}{\prod_{i \in \{a,b\}} \Gamma(\alpha_i)} \int_{\Theta_a} \int_{\Theta_b} \int_{\Theta_{\rtimes}} \prod_{i \in \{a,b,\rtimes\}} \prod_{j \in \{a,b\}} \vec{\theta}_i[j]^{n_{i \to j} + \alpha_j - 1}$$

$$= \frac{\Gamma(\sum_{i \in \{a,b\}} \alpha_i)}{\prod_{i \in \{a,b\}} \Gamma(\alpha_i)} \frac{\prod_{i \in \{a,b,\rtimes\}} \left(\prod_{j \in \{a,b\}} \Gamma(n_{i \to j} + \alpha_j) \right)}{\prod_{i \in \{a,b,\rtimes\}} \Gamma(\sum_{j \in \{a,b\}} n_{i \to j} + \alpha_j)}$$

Problem 6: Let us assume that the parameters of the Dirichlet distribution are $\vec{\alpha} = (1,1)$. Using your solution to the previous problem, compute the marginal probability of the string *aabab*. The formula that you compute should contain the Gamma function $\Gamma(\cdot)$. Using the properties of the Gamma function discussed in class (i.e., it's relationship to the factorial) or an online calculator, compute a numerical value for this expression.

Answer 6: Using the equation found in Answer 5 and an online calculator, the following holds:

$$Pr(S|\vec{\alpha}) = \frac{\Gamma(\sum_{i \in \{a,b\}} \alpha_i)}{\prod_{i \in \{a,b\}} \Gamma(\alpha_i)} \frac{\prod_{i \in \{a,b,\rtimes\}} \prod_{j \in \{a,b\}} \Gamma(n_{i \to j} + \alpha_j)}{\Gamma(\prod_{i \in \{a,b,\rtimes\}} \sum_{j \in \{a,b\}} n_{i \to j} + \alpha_j)}$$

$$Pr(aabab|(1,1)) = \frac{\Gamma(2)\Gamma(1+1)\Gamma(1+1)\Gamma(1+2)\Gamma(1+1)\Gamma(1)\Gamma(1)}{\Gamma(1)\Gamma(1)\Gamma(1+1+1)\Gamma(1+1+1)\Gamma(2+1+1+1)}$$

$$= \frac{\Gamma(2)\Gamma(2)\Gamma(2)\Gamma(3)\Gamma(2)\Gamma(1)\Gamma(1)}{\Gamma(1)\Gamma(1)\Gamma(3)\Gamma(3)\Gamma(5)}$$

$$= \frac{\Gamma(2)^4 \Gamma(3)}{\Gamma(3)^2 \Gamma(5)}$$

$$= \frac{1^4 * 2}{2^2 * 24}$$

$$= \frac{2}{96}$$

$$= \frac{1}{48}$$

$$= .020833$$

Problem 7: Suppose that we have observed a sentence $S = W^{(1)}, \ldots, W^{(k)}$. Find an expression for the posterior distribution over the model parameters, $Pr(\vec{\theta_a}, \vec{\theta_b}, \vec{\theta_{\rtimes}} \mid S, \vec{\alpha})$.

Hint: Use the joint probability that you computed in Problem 4 and Bayes' rule. The solution should be analogous to the posterior probability for the bag of words model.

Answer 7: According to conditional Bayes' Rule, the following math holds true:

$$Pr(\vec{\theta_a}, \vec{\theta_b}, \vec{\theta_{\bowtie}} | S, \vec{\alpha}) = Pr(S | \vec{\theta_a}, \vec{\theta_b}, \vec{\theta_{\bowtie}}, \vec{\alpha})$$

$$= \frac{Pr(S | \vec{\theta_a}, \vec{\theta_b}, \vec{\theta_{\bowtie}}, \vec{\alpha}) Pr(\vec{\theta_a}, \vec{\theta_b}, \vec{\theta_{\bowtie}} | \vec{\alpha})}{Pr(S | \alpha)}$$

$$= \frac{Pr(S | \vec{\theta_a}, \vec{\theta_b}, \vec{\theta_{\bowtie}}) Pr(\vec{\theta_a}, \vec{\theta_b}, \vec{\theta_{\bowtie}} | \vec{\alpha})}{Pr(S | \alpha)}$$

$$Pr(S | \vec{\theta_a}, \vec{\theta_b}, \vec{\theta_{\bowtie}}) = \frac{Pr(S, \vec{\theta_a}, \vec{\theta_b}, \vec{\theta_{\bowtie}} | \vec{\alpha})}{Pr(\vec{\theta_a}, \vec{\theta_b}, \vec{\theta_{\bowtie}} | \vec{\alpha})}$$

Therefore, $Pr(\vec{\theta_a}, \vec{\theta_b}, \vec{\theta_{\bowtie}} | S, \vec{\alpha}) = \dfrac{Pr(S, \vec{\theta_a}, \vec{\theta_b}, \vec{\theta_{\bowtie}} | \vec{\alpha})}{Pr(S | \alpha)}$

$$= \frac{\prod_{i \in \{a,b,\bowtie\}} \Gamma(\sum_{j \in \{a,b\}} n_{i \to j} + \alpha_j)}{\prod_{i \in \{a,b,\bowtie\}} (\prod_{j \in \{a,b\}} \Gamma(n_{i \to j} + \alpha_j))} * \prod_{i \in \{a,b,\bowtie\}} \prod_{j \in \{a,b\}} \vec{\theta_i}[j]^{n_{i \to j} + \alpha_j - 1}$$

Problem 8: Using your formula for the posterior distribution and setting $\vec{\alpha} = (1, 1)$, calculate the posterior distribution over parameters given the sentence *aabab*. There should be an easy way to interpret the posterior distribution, and how it was derived from the observed sentence. What is this interpretation?

Answer 8: Using the formula obtained in Answer 7 and provided values, the following holds true:

$$Pr(\vec{\theta_a}, \vec{\theta_b}, \vec{\theta_{\bowtie}} | S, \vec{\alpha}) = \frac{\prod_{i \in \{a,b,\bowtie\}} \Gamma(\sum_{j \in \{a,b\}} n_{i \to j} + \alpha_j)}{\prod_{i \in \{a,b,\bowtie\}} (\prod_{j \in \{a,b\}} \Gamma(n_{i \to j} + \alpha_j))} * \prod_{i \in \{a,b,\bowtie\}} \prod_{j \in \{a,b\}} \vec{\theta_i}[j]^{n_{i \to j} + \alpha_j - 1}$$

$$Pr(\vec{\theta_a}, \vec{\theta_b}, \vec{\theta_{\bowtie}} | aabab, (1,1)) = \frac{\Gamma(1 + 1 + 1)\Gamma(1 + 1 + 1)\Gamma(1 + 2 + 1 + 1)}{\Gamma(1 + 1)^3 \Gamma(1 + 2) \Gamma(1)^2} * (.5 * .7 * .3^2 * .2)$$

$$= \frac{\Gamma(3)^2 \Gamma(5)}{\Gamma(2)^3 \Gamma(3) \Gamma(1)^2} * .0063$$

$$= \frac{2^2 * 24}{1^3 * 2 * 1^2} * .0063$$

$$= \frac{4 * 24}{1 * 2 * 1} * .0063$$

$$= \frac{96}{2} * .0063$$

$$= 48 * .0063$$

$$= .3024$$

The posterior distribution is a conditional probability, given background information in the form of a prior distribution. This is consistent with Answer 7, as Conditional Bayes' Rule (used to determine the answer) yields a posterior distribution.

Problem 9: Consider the language $L = \{a^* b a^*\}$, that is, the language consisting of some number of a's, followed by a single b, followed by some number of a's. Show that this language is not strictly 2-local.

Hint: use n-Local Suffix Substitution Closure (n-SSC).

Answer 9: According to n-Local Suffix Substitution Closure (n-SSC), if all pivots of length $n - 1$ and all u and v that form a correct sentence in the language L around the pivot can be combined around the pivot such that u is before and v is after the pivot x, and only u, x, and v compose the sentence, then a language

is n-local. Using this logic, if a sentence uxv exists, and a sentence u'xv' exists, but uxv' does not exist in the language L, then the language is not n-local (where the length of x is $n-1$).
The following values of u, x (of length $n-1$), and v form sentences in $L = \{a^*ba^*\}$:

u	x	v	sentence
$\rtimes a$	a	$ba\ltimes$	$aaba$
$\rtimes ab$	a	$a\ltimes$	$abaa$

However, if you switch the combinations of u and v, you get the following answers, neither of which form sentences in $L = \{a^*ba^*\}$.

u	x	v	sentence
$\rtimes ab$	a	$ba\ltimes$	$ababa$
$\rtimes a$	a	a	aaa

As such, $L = \{a^*ba^*\}$ is not strictly 2-local.

Problem 10: Consider the language $L = \{a^n b^m c^n d^m \mid n, m \in \mathbb{N}\}$, that is, the language consisting of n a's followed by m b's followed by n c's followed by m d's where n and m are natural numbers. Show that this language is not strictly 2-local.

Hint: use the same property as in the problem above.

Answer 10: Using the same logic as in Answer 9, the following holds:
The following combinations of u, x, and v (where x is of length $n-1$ or 1) are in the language $L = \{a^n b^m c^n d^m \mid n, m \in \mathbb{N}\}$:

u	x	v	sentence
$\rtimes ab$	b	$cdd\ltimes$	$abbcdd$
$\rtimes a$	b	$bcdd\ltimes$	$abbcdd$

However, if you switch the combinations of u and v, you get the following answers, neither of which form sentences in $L = \{a^n b^m c^n d^m \mid n, m \in \mathbb{N}\}$.

u	x	v	sentence
$\rtimes a$	b	$cdd\ltimes$	$abcdd$
$\rtimes ab$	b	$bcdd\ltimes$	$abbbcdd$

As such, $L = \{a^n b^m c^n d^m \mid n, m \in \mathbb{N}\}$ is not strictly 2-local.

Problem 11: Show that the language $L = \{a^n b^m c^n d^m \mid n, m \in \mathbb{N}\}$ is not strictly k-local, for any value of k.

Answer 11: Using the same logic as in Answer 9 and Answer 10, the following holds:
The following combinations of u, x, and v (where x is of length n-1 or k-1) are in the language $L = \{a^n b^m c^n d^m \mid n, m \in \mathbb{N}\}$:

u	x	v	sentence
$\rtimes a^n b$	b^{k-1}	$b^{m-k} c^n d^m \ltimes$	$a^n b^m c^n d^m$
$\rtimes a^n$	b^{k-1}	$b^{m-k+1} c^n d^m \ltimes$	$a^n b^m c^n d^m$

However, if you switch the combinations of u and v, you get the following, neither of which are sentences in $L = \{a^n b^m c^n d^m \mid n, m \in \mathbb{N}\}$:

u	x	v	sentence
$\rtimes a^n$	b^{k-1}	$b^{m-k}c^n d^m \ltimes$	$a^n b^{m-1} c^n d^m$
$\rtimes a^n b$	b^{k-1}	$b^{m-k+1}c^n d^m \ltimes$	$a^n b^{m+1} c^n d^m$

As such, $L = \{a^n b^m c^n d^m \mid n, m \in \mathbb{N}\}$: is not strictly k-local for any value of k.

Problem 12: In class we proved that $k-\mathrm{SSC}(L) \implies L \in \mathrm{SL}_k$. In other words, if a language satisfies k-Local Suffix Substitution Closure, then it is k-strictly local.

Use this theorem to prove that k-strictly local languages are closed under intersection. More precisely, prove that if $L_1 \in \mathrm{SL}_k$ and $L_2 \in \mathrm{SL}_k$, then $L_1 \cap L_2 \in \mathrm{SL}_k$.

Answer 12: Below, it will be shown that k-strictly local languages are closed under intersection if they satisfy k-local suffix substitution closure (in other words, that $(L_1 \in \mathrm{SL}_k) \cap (L_2 \in \mathrm{SL}_k) \in \mathrm{SL}_k$). The details of k-local SSC are described in Answer 9.

Proof. First, we will assume that L_1 and L_2 are both in SL_k.
Thus, according to the definition of a strictly n-local language we know that there exist two grammars ($G_1 = <V, T_1>$ and $G_2 = <V, T_2>$, respectively) such that T_1 and T_2 are subsets of $F_k(\rtimes \cdot v^* \cdot \ltimes)$.
Let L equal the language formed as follows: $G = <V, T_1 \cap T_2>$. In other words, sentences in language L can be generated by factors in the intersection of T_1 and T_2.
Note that the intersection of T_1 and T_2 contains factors in either T_1 or T_2 or both, and not any others. As such, it is trivial that sentences generated with such factors are in the intersection of the two languages that these factors map to, and not any others.
Thus, we understand how $L = L_1 \cap L_2$ is formed, and can see how it satisfies the definition of a strictly k-local grammar.
This point will now be further emphasized, using the theorem that $k-\mathrm{SSC}(L) \implies L \in \mathrm{SL}_k$.
Let there be two random, arbitrary strings s_1 and s_2 such that $s_1 = u_1 x v_1 \in L$ and $s_2 = u_2 x v_2 \in L$, where x is a pivot of length $k-1$ (as it was in Answers 9-11).
By the definition of a substring, we know that $u_1 x$ and $u_2 x$ are substrings of s_1 and s_2, respectively. As such, it is trivial that $F_k(u_1 x) \in T_G$ and $F_k(u_2 x) \in T_G$, where $T_G = T_1 \cap T_2$.
By permuting the prior and post pivot segments, the following holds, thus proving that k-strictly local languages are closed under intersection:

$$F_k(\rtimes u_1 x v_2 \ltimes) = F_k(\rtimes u_1 x) \cup F_k(x v_2 \ltimes)$$
$$\implies F_k(\rtimes u_1 x v_2 \ltimes) \in T_G$$
$$\implies u_1 x v_2 \in L$$
$$\implies K - SSC(L) \implies L \in SL_k$$

$\square$