Panini Linguistics Olympiad 2018

Seniors (Class IX to XII)

Round 1, 7th January 2018

Question Booklet

Maximum Marks: 100   **Total Time: 4 hours**

Instructions:

1. There are 7 printed pages in this question booklet, including this page. If your booklet has less than 7 printed pages, report to the invigilator immediately.
2. There are 5 problems of 20 marks each. The problems have many sub-parts, all of which must be answered to receive full credit for the problem.
3. The Answer Booklet provided to you has specific space for writing down the objective solutions of each problem. **You MUST write down the precise answers or solutions to the problems in the space provided.** The assignments for which at least a part of the answer needs to be written down in the answer booklet have been specifically indicated.
4. The details and explanations of your answers and the rules of the language should be written in separate sheets.
5. While explaining your answers in separate sheets, you need not rewrite the solutions that you have already provided in the Answer Booklet.
6. Write down your explanations to each problem on a new sheet or sheets of paper. On each sheet, indicate the **number of the problem**, your **roll number**, and your **name**. Otherwise, your work may be mislaid or misattributed.
7. Do not copy the statements of the problem.
8. **All answers must be well-argued. Even a perfectly correct answer will receive a low score unless accompanied by an explanation.**
9. Each problem has been thoroughly checked for clarity, accuracy and solvability. Some problems are more difficult than others, but all can be solved using ordinary reasoning and some basic analytical skills. You don't need to have prior knowledge of linguistics or these languages in order to solve them.
10. The question paper has been designed to ensure that very few people will solve all these problems completely in the time allotted. Don't be discouraged if you don't finish everything.
11. Use of calculators, mobile phones and any other electronic devices is strictly prohibited. No books, notebooks or other printed materials can be consulted during the contest.

Good luck!

PLO 2018 Committees and Organizing Institutions

Problem Committee

Chair:

Monojit Choudhury, Microsoft Research India

Editor:

Saujas Vaduguru, IIIT Hyderabad

Members:

Abhishek Dedhe, University of Rochester

Chinmaya Kausik, IISc Bangalore

Ritesh Kumar, B. R. Ambedkar University, Agra

Bornini Lahiri, Jadavpur University

Vasundhara Rakesh, IIT Kanpur

Sesh Sadasivam, University of Michigan

Aalok Sathe, University of Richmond

Yash Sinha, University of Chicago

Anshuman Swain, University of Maryland

Jury

Chairs:

Monojit Choudhury, Microsoft Research India

Dipti Misra Sharma, IIIT Hyderabad

Members:

Kalika Bali, Microsoft Research India

Anshul Bawa, Microsoft Research India

Sandipan Dandapat, Microsoft India

Radhika Mamidi, IIIT Hyderabad

Adithya Pratapa, Microsoft Research India

Sunayana Sitaram, Microsoft Research India

Nirmal Surange, IIIT Hyderabad

Saujas Vaduguru, IIIT Hyderabad

Organizing Committee

Chairs:

Dipti Misra Sharma, IIIT Hyderabad

Manish Srivastava, IIIT Hyderabad

Members:

Kalika Bali, Microsoft Research

Monojit Choudhury, Microsoft Research

Prathyusha Danda, IIIT Hyderabad

Tirthankar Dasgupta, TCS Innovation Labs

Organizing Committee (contd.)

Asif Ekbal, IIT Patna

Girish Nath Jha, JNU, New Delhi

Preethi Jyothi, IIT Bombay

Mitesh Khapra, IIT Madras

Ritesh Kumar, B.R. Ambedkar University, Agra

Shakuntala Mahanta, IIT Guwahati

Radhika Mamidi, IIIT Hyderabad

Pruthwik Mishra, IIIT Hyderabad

Sibansu Mukhopadhyay, SNLTR, Kolkata

Soma Paul, IIIT Hyderabad

Avinash Pandey, Mumbai University

Ashutosh Tripathi, Krishna Public School Raipur

Saumitra Yadav, IIIT Hyderabad

Organizing Institutions

National Coordinator: IIIT Hyderabad

Regional Coordinators:

B. R. Ambedkar University, Agra

IIT Guwahati

IIT Madras

IIT Patna

Jawaharlal Nehru University, Delhi

Microsoft Research India, Bangalore

Mumbai University

SNLTR, Kolkata

Test Centres:

SNLTR, Kolkata

Krishna Public School, Raipur

Kendriya Hindi Sansthan, Agra

IIT Madras

SP College, Pune

Microsoft Research India, Bangalore

IIIT Hyderabad

IIT Patna

IIT Bombay

Delhi Public School, Udaipur

Center for Sanskrit Studies, JNU, Delhi

1. Sixteen Japanese Phrases (20 marks)

Problem designed by Abhishek Dedhe

Given below is a list of 16 Japanese phrases, followed by their English translations in random order.

niko tamago, gohon kawa, hachihon gen, goko ringo, kyuhiki inu, gomai pureto,
nihiki neko, rokumai enberopu, gohiki neko, rokumai pureto, nihon hashi, kyuhon kawa,
kyumai kitte, hachihiki ari, kyuko tama, rokuko tama

two eggs, five apples, five plates, two wildcats, nine balls, five rivers, nine rivers, nine dogs,
six envelopes, eight ants, nine postage stamps, two chopsticks, eight strings,
five wildcats, six plates, six balls

Assignment 1 [Answer Booklet]: Match the translations.

Assignment 2 [Answer Booklet]: Given your current knowledge and that *tiger* translates to *tora* in Japanese, which of the following would be valid and why?

nihiki tora, niko tora, nihon tora, nimai tora

Assignment 3 [Answer Booklet]: Actually, none of the above four phrases are valid. Instead, you might expect to see *nito tora*. You might also see *nito zo* (*zo = elephant*), and *nito raion* (*raion = lion*).

Which of the following are correct/valid Japanese phrases? Provide the English translations for the correct ones. Note: *yajirushi = arrow*, *hachi = bee*.

*hachimai hashi, rokuhon yajirushi, gohiki zo, niko ringo, hachiko hashi, goto neko*

Notes: Japanese, also known as *Nihongo*, is an East Asian language spoken by about 126 million people, primarily in Japan, where it is the official language and national language. It is a member of the Japonic (or Japanese-Ryukyuan) language family, and its relation to other languages, such as Korean, is debated. Japanese has been grouped with language families such as Ainu and Austroasiatic, but none of these proposals has gained widespread acceptance. Japanese has no genetic relationship with Chinese, but it makes extensive use of Chinese characters, or *kanji*, in its writing system, and a large portion of its vocabulary is borrowed from Chinese.

2. Word Embeddings (20 marks)

Problem designed by Monojit Choudhury

Computers can "compute" mathematical functions that take numbers as input and give numbers as output. Therefore, in order to make computers understand and process human language, it is useful to represent words in terms of numbers. Computer scientists have come up with sophisticated techniques that encode words as "vectors", i.e., a sequence of numbers. These are called *word embeddings.* It might be easier to think of word embedding as a mapping of words to points in an n-dimensional space. Let us assume that n = 2, i.e., words are mapped to points in a two-dimensional plane.
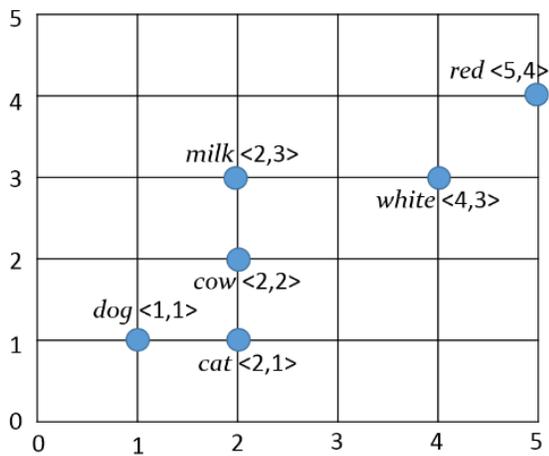
The diagram in the left shows hypothetical word embeddings for 6 English words. As you can see in this example, words which have similar meanings or are used in similar context have word embeddings which are close to each other, i.e., the points are closer in space. Note: The embeddings in this example are NOT related to the problem below.

Here are some English words in alphabetical order:

boy, first, girl, grammar, language, literature, man, mathematics, mathematician, number, one, position, queen, second, time, two, woman

Shown below, in no particular order, are the translations of these words in the Santali language written in the Ol Chiki script, and their (the English words') embeddings.

#	Santali Translation	English Embedding
A	ᱢ (Ol Chiki script)	<3,3>
B	ᱢ (Ol Chiki script)	<3,2>
C	ᱢ (Ol Chiki script)	<5,4>
D	ᱢ (Ol Chiki script)	<7,2>
E	ᱢ (Ol Chiki script)	<7,4>
F	ᱢ (Ol Chiki script)	<8,2>
G	ᱢ (Ol Chiki script)	<8,4>
H	ᱢ (Ol Chiki script)	<9,3>
I	ᱢ (Ol Chiki script)	<10,6>

#	Santali Translation	English Embedding
J	ᱢ (Ol Chiki script)	<8,13>
K	ᱢ (Ol Chiki script)	<4,7>
L	ᱢ (Ol Chiki script)	<5,6>
M	ᱢ (Ol Chiki script)	<3,8>
N	ᱢ (Ol Chiki script)	<9,10>
O	ᱢ (Ol Chiki script)	<11,11>
P	ᱢ (Ol Chiki script)	<8,12>
Q	ᱢ (Ol Chiki script)	<10,13>

Assignment 1 [Answer Booklet]: Match the English words to their Santali translations – A to Q. No explanation needed.

Assignment 2 [Answer Booklet]: Can you guess the embedding of the words: *king, linguist, princess*? Explain your answer.

Assignment 3 [Answer Booklet]: Give the English translation of the following Santali word.

ᱢ (Ol Chiki script)

Notes: Santali is a language in the Munda subfamily of Austroasiatic languages. It is spoken by around 6.2 million people in India, Bangladesh, Bhutan and Nepal. Most of its speakers live in India, in the states of Jharkhand, Bihar, Odisha, Tripura, Mizoram, Assam and West Bengal. Since Santali did not have a script, it was written using the Roman script or the Eastern Nagari (Bangla) script. However, none of the existing scripts were able to phonetically represent the Santali language. This resulted in the invention of a new script called *Ol Chiki* by Pandit Raghunath Murmu in 1925. He is popularly known as Guru Gomke among the Santals.

3. Kurmali (20 marks)

Problem designed by Bornini Lahiri

Here are sentences in Kurmali and their respective English translations.

1.	*moi bat khau*	*I ate rice*
2.	*moira bat khayonu*	*We ate rice*
3.	*toi choani ke dekhliae*	*You (Masculine) saw a girl*
4.	*toira choara ke dekhlini*	*You all saw the boys.*
5.	*choara choani ke dekhaone*	*Boys saw a girl.*
6.	*choara ghar dekhaonu*	*Boys saw the house.*
7.	*kokur kha*	*Dog ate.*
8.	*kokura khayon*	*Dogs ate.*
9.	*toi kokur ke dekhleou*	*You (feminine) saw a dog.*
10.	*Radha Salman ke dekhuoi*	*Radha saw Salman.*
11.	*choanira bhat khaenu*	*Girls ate rice.*
12.	*Salman Radha ke putuli deaoe*	*Salman gave a doll to Radha.*

Assignment 1 [Answer Booklet]**:** Translate the following sentences in Kurmali.

1. I saw Radha.
2. I gave Salman rice.
3. Girls ate the rice.
4. A girl gave Radha dolls.
5. Radha saw the boys.
6. We saw a dog.

Assignment 2: Explain how the grammar of the language works.

Notes: (1) Kurmali is a language spoken in some of the districts of West Bengal, Jharkhand, Orissa and Madhya Pradesh. It is believed to be a mixed language. (2) The data has been modified slightly for ease of solvability.

---

4. Imya Imyevna Rossiyov (20 marks)

Problem designed by Saujas Vaduguru

You are studying a Russian language and culture course in college, and as part of an assignment, you have to figure out how the Russian family naming system works. You are given the scheme of a family tree, and the statements by people where they name some people to whom they are related. These relatives may be their parent, child, or sibling. The statements include names of some people who do not fit in the given scheme for the family tree. In the figure, an arrow goes from a parent to a child. Rounded nodes are for female names and rectangular ones are for male names.

The following statements (naming relatives) are from members represented in the family tree — A through L. But, the statements themselves may contain names of people who are not represented in the family tree, i.e., the statements may name people who are not one of A-L.

Note that (1) *Vladimir* is male, and *Yekaterina* is female. (2) There are two different *Olga*s in the family tree. (3) The first names are followed by corresponding nicknames in brackets.
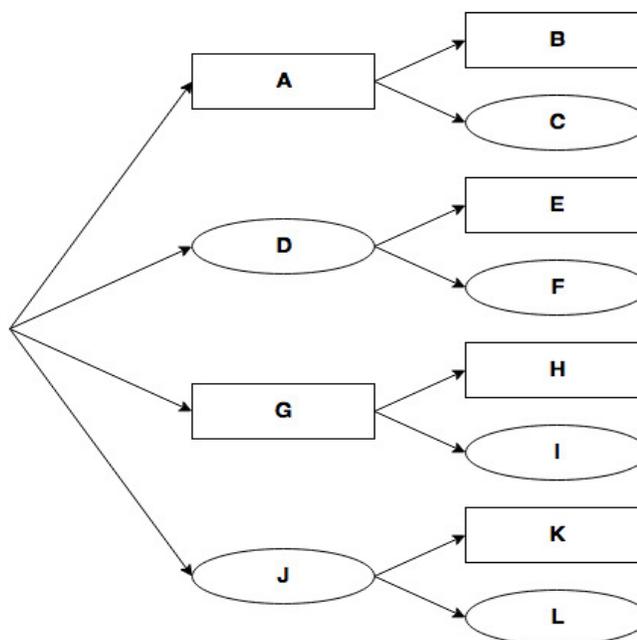
Vladimir (Vova): *Vanyechka, Andriy Vladimirovich Ivanov, Olga Vladimirovna Ivanova*
Viktor (Vitya): *Alexandr Romanovich Smirnov, Natalya*
Valeriya (Lera): *Anna Vladimirovna Ivanova, Sergey Pavelovich Popov*
Anna (Anya): *Lyonyechka, Lerochka, Olga Vladimirovna Ivanova, Andriy Vladimirovich Ivanov*
Olga (Olya): *Vityechka, Natashochka, Anna Vladimirovna Ivanova, Andriy Vladimirovich Ivanov, Vladimir Vladmirovich Ivanov*
Andriy: *Vladimir, Olechka, Boryechka*
Leonid (Lyonya): *Anna Vladimirovna Ivanova, Valeriya Sergeyevna Popova*
Nataliya (Natalya): *Viktor Alexandrovich Smirnov*
Olga (Olya): *Boris, Andriy Vladimirovich Ivanov*
Ivan (Vanya): *Yelena, Yekaterina Andriyevna Mikhailova*
Boris (Borya): *Andriy Vladimirovich Ivanov, Olga Andriyevna Ivanova, Alexandra Georgyevna Vasileva*
Yelena (Lena): *Yekaterina Andriyevna Mikhailova, Vladimir Vladmirovich Ivanov*

Assignment 1 [Answer Booklet]: Fill in the full, three-part names of the family members A-L in the family tree. If there are multiple possible solutions, write only one.

Assignment 2: Explain how the Russian naming system works.

Assignment 3 [Answer Booklet]**:** There is a grammatical error in the title of the problem. What is it?

Notes: Russian is an East Slavic language and an official language in Russia, Belarus, Kazakhstan, Kyrgyzstan and many minor or unrecognized territories. It is the most geographically widespread language of Eurasia. It is also the largest native language in Europe, with 144 million native speakers in Russia, Ukraine and Belarus.

5. Washo (20 marks)

Problem designed by Yash Sinha

Consider the following sentences from Washo, and their English translations.

(1) *teliwhu basa' basa'ige da'mo'mo' sili*   The man gave to the woman the book that he wrote.

(2) *basa' digyisge teliwhu disili*   The man gave to me the book that I saw.

(3) *da'mo'mo' basa' basa'isge dipigelyayi*   I returned to the woman the book that she wrote.

(4) *basa' umbasa'ige teliwhu umsili*   You gave to the man the book that you wrote.

(5) *teliwhu basa' umsilisge dipigelyayi*   I returned to the man the book that he gave to you.

(6) *basa' umgyige da'mo'mo' umsili*   You gave to the woman the book that you saw.

(7) *teliwhu da'mo'mo' basa' silige gyi*   The man saw the book that he gave to the woman.

Assignment 1 [Answer Booklet]**:** Translate the following sentences into English. If there are multiple translations possible, give all of them.

a. *da'mo'mo' basa' umsilige basa'i*

b. *teliwhu basa' disilisge digyi*

c. *da'mo'mo' basa' disilisge dipigelyayi*

Assignment 2 [Answer Booklet]**:** Translate the following sentences into Washo.

a. The man gave to me the book that I returned to him.

b. You saw the book that you gave the woman.

Assignment 3: Explain how the language works by describing the order in which words are used in a sentence, the translations for different verbs and nouns, the way in which verbs and nouns get modified and any other things that you find interesting about the language.

Notes: (1) Washo is an endangered Native American language isolate spoken by the Washo on the California–Nevada border in the drainages of the Truckee and Carson Rivers, especially around Lake Tahoe. While there are only 20 elderly native speakers of Washo, since 1994 there has been a small immersion school that has produced many moderately fluent younger speakers. This problem is based on data from Emily Hanink's field work. (2) Some of the words in this problem have been altered slightly for ease of solvability; for the same reason, diacritics have been omitted.

---END OF QUESTION BOOKLET--