

10-11 классы

Задача 1 (30 баллов)

Для кодирования грамматической информации в компьютерной лингвистике применяются специальные системы помет, которые представляют лингвистические сведения в стандартизированной форме. Одна из таких систем, предназначенных, в частности, для русского языка, носит название MULTEXT-East. Даны некоторые пометы, используемые в этой системе для разметки существительных; к каждой помете приводится по примеру:

Ncmpru — участников

Ncfsnn — поддержка

Npmsgy — Владимира

Ncnsin — удовольствием

Ncmpln — пределах

Ncfply — уборщицах

Ncmsan — режим

Npfsay — Юлию

Ncfpdn — группам

1.1. Схематично опишите, как устроена система помет существительных в системе MULTEXT-East.

1.2. Некоторые из девяти слов, данных в условии, могли бы иметь в системе системе MULTEXT-East и другие пометы. Укажите эти слова и приведите эти пометы.

1.3. Как будут выглядеть в системе MULTEXT-East пометы существительных *Иванову, животные, гордячкой*? Если в каком-то случае возможны несколько вариантов, укажите их все.

Решение и критерии:

3.1.

Каждая помета имеет постоянную длину и состоит из 6 символов. За каждой позицией закреплён определённый набор символов с определённым значением:

1. N — всегда в начале пометы существительного (впрочем, пометы других частей речи нам и не даны)
2. c / p — нарицательное / собственное (2 балла). За ответ «с большой буквы / с маленькой буквы» балл не ставился, так как это не грамматическая информация
3. m / f / n — род (2 балла)
4. s / p — число (2 балла)
5. n / g / d / a / i / l — падеж (2 балла)
6. y / n — одушевлённость (2 балла)

Для получения баллов необходимо было **точно** указать, на какой позиции используется какой из наборов символов. Кроме того, не засчитывались как правильные нерасчленённые описания (например, когда говорилось, что за падеж отвечают два последних символа). Обозначения категорий рода, числа и падежа не вызвали проблем у участников, однако оказалось, что многие

не помнят термины «одушевлённость», «собственные» и «нарицательные». Такие решения оценивались полным (2 из 2) или частичным баллом (1 из 2) в зависимости от того, насколько успешно школьник сумел объяснить, что он имеет в виду, пусть даже не прибегая к термину.

Особую сложность вызвало определение падежа и одушевлённости. Кратко напомним несколько основных фактов, касающихся этих грамматических категорий:

- Одушевлённость в русской грамматике определяется строго формально, а не по смыслу; одушевлённым является такое существительное, у которого во множественном числе вин. = род. (*вижу участников = выступления участников ≠ участники выступают*), а неодушевлённым — такое, у которого во множественном числе вин. = им. (*вижу группы = группы выступают ≠ выступления групп*); в единственном числе этот критерий применим к существительным мужского рода, за исключением слов I склонения (*папа, воевода*); кроме того, одушевлённость можно определять также по согласуемому слову (прилагательному / местоимению) во множественном числе и в единственном числе мужского рода. Из всего этого следует, что слово *группа* не может быть одушевлённым, даже если группа, о которой идёт речь, состоит из живых людей (мы всё равно скажем *вижу студенческие группы*, а не **вижу студенческих групп*), а слово *участник* не может быть неодушевлённым (*вижу участника/участников*, а не **вижу участник/участники*) — соответственно, вторые варианты разбора в этих случаях предлагать не следовало, баллы за это снижались.
- Из-за наличия категории одушевлённости в русском языке не всегда просто различить именительный, родительный и винительный падежи. Практический способ решения этой проблемы таков: нужно подставить на место определяемого слова любое слово I склонения в единственном числе (*кошка, страна, папа, дисциплина, ...*), поскольку такие слова различают эти три формы. Например, форма *режим* может употребляться в тех же контекстах, что формы *дисциплина* и *дисциплину* (*царит строгий режим / строгая дисциплина, соблюдать режим / дисциплину*), но не *дисциплины*, а значит, это может быть только именительный или винительный падеж единственного числа.
- ◆ Как справедливо отметили многие участники олимпиады, в основе системы помет лежат английские слова. При этом для решения задачи знать английский язык вовсе необязательно: например, число могло бы обозначаться не буквами s и p, а любыми другими символами (например, q и w), но мы могли бы точно так же расписать все примеры на две группы и понять, что раз знак q стоит на 4-м месте в пометах слов *поддержка, Владимира, удовольствием, режим, Юлию*, а знак w — на 4-м месте в пометах слов *участников, пределах, уборщицах, группам*, то q обозначает единственное число, а w — множественное. Многие участники пытались приводить английские (а иногда — немецкие, французские или латинские) слова, от которых образованы обозначения в помете, но мало кому удавалось сделать это правильно. Разумеется, ошибки в словах иностранных языков никак не наказывались на олимпиаде по русскому языку, но всё же приведём правильный перечень английских грамматических терминов, легших в основу этой системы — они могут пригодиться при чтении лингвистических научных работ:
 1. N — noun ‘существительное’ (а не name ‘имя’)
 2. c / p — common ‘нарицательное’ / proper ‘собственное’ (а не personal ‘личное’)
 3. m / f / n — masculine ‘мужской (о грамматическом роде)’ (а не male ‘мужской (о поле)’, man ‘мужчина’) / feminine ‘женский (о грамматическом роде)’ (а не female ‘женский (о поле)’) / neuter ‘средний (о грамматическом роде)’ (а не neutral ‘нейтральный’, none ‘никакой’, no ‘нет’)
 4. s / p — singular ‘единственное’ / plural ‘множественное’
 5. n / g / d / a / i / l — nominative ‘именительный’ / genitive ‘родительный’ / dative ‘дательный’ / accusative ‘винительный’ / instrumental ‘творительный’ (букв.

инструментальный) / locative ‘предложный (букв. местный, т. к. предложный падеж часто выражает значение места)’

- б. у / п — yes ‘да’ / no ‘нет’ (а ‘одушевлённый’ / ‘неодушевлённый’ на английском языке будет animate / inanimate).

3.2. В этом задании требовалось предложить пометы для омонимичных форм, используя имеющуюся систему (а не усовершенствовать систему помет, например, добавив в неё тип склонения, как предлагали некоторые участники). Заметить все возможные случаи омонимии оказалось непросто. Для получения полного балла было необходимо указать пять основных случаев:

- **Ncmprgy** — участников (род., а не вин.)
- **Npmsay** — Владимира (вин. мужского имени, а не род.)
- **Npmsgn** — Владимира (род. неодушевлённого названия города, а не одушевлённого мужского имени)
- **Ncmsnn** — режим (им., а не вин.)
- **Npmsdy** — Юлию (дат. мужского имени *Юлий*, а не вин. женского имени *Юлия*)

2 балла × 5 = **max 10 баллов**; если в помете допущена ошибка хотя бы на 1 букву, ставится **0 баллов**. Давались дополнительные **2 балла** за редкое женское имя *Владими́ра* (**Npfsny**), но его упоминание не считалось обязательным.

3.3.

- *Иванову* — **Npmsdy** (дат. мужской фамилии *Иванов*);
- *Иванову* — **Npfsay** (вин. женской фамилии *Иванова*);
- *Иванову* — **Npnsdn** (дат. названия города *Иваново*); засчитывался также вариант **Npmsdn**, если указывалось, что это дат. падеж топонима *Иванов*. Этот вариант оказался даже более популярным, чем дат. от *Иваново*. По крайней мере один населённый пункт с названием *Иванов* действительно существует (в Винницкой области Украины), хотя едва ли он известен многим участникам олимпиады; вероятно, многие школьники сочли, что слово *Иваново* не изменяется по падежам, поэтому этот вариант вспомнили далеко не все;
- *животные* — **Ncnpny**; это слово является одушевлённым (об одушевлённости см. выше), так как вин. мн. у него совпадает с род. мн. (*вижу животных* = *детёныши животных* ≠ *животные бегают*). Соответственно, последним символом пометы должно быть у, а разбора с винительным падежом эта форма не имеет.
- *гордячкой* — **Ncfsiy**; это слово обозначает гордую женщину и является одушевлённым (*вижу гордячек* = *лица гордячек* ≠ *гордячки идут*); многие участники предлагали также неодушевлённый разбор, однако такого неодушевлённого слова в русском языке нет: по всей вероятности, они перепутали слова *гордячка* и *горячка* (а чрезмерная гордость называется *гордыней*, а не *гордячкой*).

2 балла × 5 = **max 10 баллов**; если в помете допущена ошибка на 1 букву, ставится **1 балл**; если допущены ошибки более чем на 1 букву, ставится **0 баллов**.

Задача 2 (35 баллов)

Некоторые глаголы русского языка (например: *осознавать*, *мёрзнуть*, *чихать*) обладают неким свойством (элементом смысла), которое делает невозможным или очень необычным их употребление в ряде конструкций (такую невозможность принято отмечать звёздочкой). В частности:

- 1) В инфинитивных предложениях: **А Петя ну осознавать свои ошибки!* **А Маша давай мёрзнуть!* Хотя возможно: *А Петя ну бежать!* *А Маша давай петь!*
- 2) В безличных предложениях с субъектом в дательном падеже и сказуемым в возвратной форме: **Мне с трудом осознавалось.* **Пете хорошо чихается на улице.* Хотя возможно: *Сегодня мне бежалось гораздо легче.* *Пете хорошо работается на новом месте.*
- 3) В настоящем времени в значении запланированного будущего: **Завтра Петя осознаёт всю тяжесть своего положения.* **На следующей неделе мёрзну по 2 часа в день.* Хотя возможно: *Завтра Маша бежит 3 км.* *На следующей неделе работаю по 8 часов в день.*
- 4) В комбинации с обстоятельствами цели (с предложениями *ради*, *для*): **Петя осознаёт это для неё.* **Она чихает ради справки.* Хотя возможно: *Он работает ради зарплаты.* *Маша поёт для него.*

2.1. Объясните, что это за свойство.

2.2. В каких случаях глаголы, обладающие этим свойством, всё же могут употребляться в указанных конструкциях?

2.3. Приведите ещё 2 глагола, обладающих рассматриваемым свойством.

2.4. В каких ещё конструкциях глаголы *осознавать*, *мёрзнуть*, *чихать* и им подобные ведут себя отлично от остальных глаголов (не обладающих рассматриваемым свойством)? Для каждой найденной вами конструкции приведите по 2 примера.

Решение и критерии:

2.1. Нельзя *осознавать*, *мёрзнуть* и *чихать* по заказу. Эти процессы происходят помимо нашей воли, мы не можем ими управлять. Это свойство ряда глаголов лингвисты называют **неконтролируемостью**. Конечно, от участников олимпиады термина не требовалось, достаточно было описать это свойство своими словами. За точный ответ давалось **15 баллов**, за недостаточно понятные объяснения — **7 баллов**. Например, 7 баллов получали решатели, которые писали, что действия, описанные глаголами из условия, нельзя *запланировать*. Запланировать их действительно нельзя, но почему?

Приведем примеры частых ответов, за которые баллов не давалось.

- 1) «Это глаголы состояния». Действительно, многие неконтролируемые глаголы являются глаголами состояния, но не все. *Чихать* — это активное действие, поэтому гипотеза про глаголы состояния не проходит. С другой стороны, не все состояния являются неконтролируемыми (напр., *медитировать*).
- 2) «Это непереходные глаголы». Неверно: *осознавать* — переходный глагол (*осознавать факт*). Другие глаголы, данные в условии, — *петь*, *бежать*, *работать* — тоже непереходные, однако, в отличие от *чихать*, *мёрзнуть* и *осознавать*, употребляются во всех приведенных конструкциях.
- 3) «Это глаголы мгновенного действия, действие происходит быстро, здесь и сейчас». Неверно: например, *мёрзнуть* можно долго, в прошлом, настоящем и будущем.

4) «Это глаголы, описывающие нерегулярные события». Тоже неверно: можно, например, *мёрзнуть каждую зиму* или *чихать каждую весну* (при аллергии на цветение деревьев).

2.2. Неконтролируемые глаголы могут употребляться в указанных конструкциях в шутку, иронично (*Пете отлично болелось дома во время контрольной*), при игре слов, в театральные ремарках. А также когда субъект намеренно вызывает состояние или же только делает вид, что выполняет неконтролируемое действие. Например: *Выйди к Пете, он уже два часа мёрзнет ради тебя*. Или: *Давай договоримся так: я чихаю, значит, кто-то идёт*. Или: *Он упал, чтобы поймать мяч*. Ещё один возможный случай — если глаголы употреблены в переносном значении. Перечислив два из приведенных случаев и (обязательно!) снабдив их примерами, можно было получить **6 баллов**, за один случай давалось **3 балла**.

Многие школьники писали, что глаголы можно употребить в настоящих конструкциях, если поменять их вид, добавить приставку или поставить перед инфинитивом другой глагол (например, нельзя *А он давай мёрзнуть*, но можно *Он начал мёрзнуть*). Это и правда так, но ведь вопрос был: *В каких случаях глаголы, обладающие этим свойством, всё же могут употребляться в указанных конструкциях?* Добавляя глагол перед инфинитивом, мы меняем конструкцию, а заменяя вид глагола или добавляя приставку, меняем даже сам глагол.

2.3. Подходят, например, глаголы *болеть, умирать, поскользнуться, споткнуться, падать, слышать, видеть, понимать, забывать, зевать, дрожать, краснеть, удивляться, испытывать (чувство)* и др. За два и больше подходящих глагола ставилось **6 баллов**, за один — **3 балла**. За глаголы *кашлять, дышать, моргать, любить, ненавидеть* и некоторые другие баллов не давалось, так как описанными действиями человек может в какой-то степени управлять. Так, врач может попросить: *Покашляйте, пожалуйста! Дышите, не дышите! Моргните!* Эти глаголы относятся к пограничному случаю, и отличить от неконтролируемых их непросто, но многим участникам олимпиады это удалось.

2.4. Неконтролируемые глаголы, как правило, не могут употребляться:

- а) в повелительном наклонении: **А ну живо мёрзни!* **Зевай давай!* и вообще в призывах к действию: **Давай дрожать!* **Пойдём видеть!*
- б) с глаголами планирования, намерения: **Он собирается мёрзнуть,* **Она решила слышать,* **Я планирую удивляться.*
- в) с глаголами долженствования, принуждения: **Ты должен (обязан) чихать,* **Ей нужно (надо) зевать,* **Здесь запрещено умирать!*
- г) с некоторыми наречиями, указывающими на качество действия: **Он хорошо дрожал, с энтузиазмом падал, усердно болел.* Но тут надо было быть аккуратным в формулировках, так как можно сказать *Он сильно мёрз,* *Он громко чихал.*

Принимались также и любые другие подходящие ответы, содержащие примеры. Придумать работающий тест на неконтролируемость — это задача для настоящего

лингвиста, поэтому уже за один верный ответ ставился максимум — **8 баллов**. Но иногда участники приводили конструкции, уже данные в условии или очень близкие к тому. Такие ответы не засчитывались.

Задача 3 (35 баллов)

В Национальном корпусе русского языка содержатся тексты общим объёмом более 500 млн слов. В нём можно осуществлять разные поисковые запросы. Ниже дана верхняя часть списка, полученная по некоторому запросу. Одно слово при этом заменено на X.

| | | | | | | | |
|---|------|-----------------|---|----|------|--------------------|---|
| 1 | 9333 | <i>известно</i> | X | 9 | 2179 | <i>естественно</i> | X |
| 2 | 9038 | <i>ясно</i> | X | 10 | 2175 | <i>интересно</i> | X |
| 3 | 6198 | <i>видно</i> | X | 11 | 1696 | <i>мало</i> | X |
| 4 | 5663 | <i>хорошо</i> | X | 12 | 1475 | <i>странно</i> | X |
| 5 | 4875 | <i>понятно</i> | X | 13 | 1284 | <i>вероятно</i> | X |
| 6 | 4382 | <i>очевидно</i> | X | 14 | 1264 | <i>похоже</i> | X |
| 7 | 3063 | <i>возможно</i> | X | 15 | 1236 | <i>несомненно</i> | X |
| 8 | 2565 | <i>равно</i> | X | | | | |

3.1. Как был сформулирован запрос?

3.2. Что за цифры приведены во втором столбце списка?

3.3. Какое слово заменено на X?

3.4. Приведите ещё 3 примера слов, которые, скорее всего, тоже оказались высоко в этом списке, и 3 примера слов, которые оказались в его нижней части или вовсе отсутствуют в нём.

3.5. Некоторые из вышеуказанных 15 слов оказались в списке по иной причине, чем остальные. Назовите эти слова и объясните причины, по которым они оказались в верхней части списка.

Решение и критерии:

Национальный корпус русского языка (НКРЯ, www.ruscorpora.ru) – большое собрание текстов в электронной форме, наделенное широкими возможностями поиска благодаря специально разработанной разметке. Многие участники по каким-то причинам сочли, что в НКРЯ есть лишь научные тексты, однако там есть тексты самых разных жанров: проза, поэзия, драматургия, техническая и научная литература, публицистика, личная переписка, блоги и др.

3.1. Запрос был сформулирован так. Выдать в порядке убывания частотности самые частотные из встретившихся в НКРЯ цепочки слов следующего вида: «прилагательное (1

балл) среднего рода (**1 балл**) в краткой форме (**1 балл**)» плюс слово *что* в именительном падеже.

Во многих работах говорилось о наречиях. Действительно, все приведенные в списке слова можно интерпретировать и как наречия (почти все русские краткие прилагательные омонимичны наречиям). Однако если бы в запросе было задано наречие, верхняя часть списка была бы другой: в нее попали бы и такие словосочетания, как *потому что, только что, так что, пока что* и др. (поскольку слова *потому, только, так, пока* имеют, помимо других, и функции наречий). Для правильного решения надо было заметить, что в результаты поиска не попали такие слова, которые нельзя интерпретировать как прилагательные. По тем же причинам не ставились в этом пункте баллы за решения, где говорилось о вводных словах или вообще не отмечалась частеречная принадлежность («искали слова, стоящие перед *что*»).

3.2. Частота встречаемости соответствующих цепочек слов в НКРЯ (**5 баллов**).

3.3. Слово *что* (**2 балла**). Другие приводимые в работах слова (*одно, всё, ли, было* и др.), если бы были включены в запрос, дали бы другие результаты: вряд ли бы настолько частотными оказались словосочетания *мало одно, похоже одно, очевидно всё, равно всё, странно ли, несомненно ли, хорошо было, равно было* и т.п. Важно отметить, что поиск производился именно с указанным в условии порядком слов (то есть словосочетания *одно хорошо, всё равно* или *было ясно* не были бы найдены).

3.4. Среди других достаточно частотных словосочетаний данного типа можно указать следующие: *важно, что; любопытно, что; удивительно, что; замечательно, что; приятно, что; жалко, что; характерно, что; обидно, что; примечательно, что; досадно, что; страшно, что; плохо, что; заметно, что; стыдно, что; немудрено, что; грустно, что; показательно, что; смешно, что; забавно, что; ужасно, что; общеизвестно, что; поразительно, что; нехорошо, что; бесспорно, что; существенно, что; отрадно, что; глупо, что* и др. (по **2 балла** за пример, максимум **6 баллов**). Среди слов, не оказавшихся в списке или оказавшихся в нем низко, надо было привести любые краткие прилагательные, обычно не стоящие перед словом *что* (например: *бело, старательно, сонно, огромно, постоянно* и др.).

3.5. Во всех данных примерах, кроме двух, слово *что* — союз, который вводит придаточное изъяснительное предложение (*Известно, что задача трудная; Хорошо, что я ее решил; Жалко, что не на полный балл* и т.п.) (**7 баллов**). Однако два прилагательных попали в верхушку результатов по иной причине: *равно* — из-за частотного выражения *все равно что* (**2 балла**); *мало* — из-за частотного выражения *мало что* (**2 балла**). (Если эти слова приводились без объяснения, ставился **1 балл**). Так получилось потому, что при поиске в НКРЯ, как видно, не учитываются знаки препинания (**2 балла**).