М.Б. Антонова
А.В. Бакулев

АКАДЕМИЧЕСКОЕ ПИСЬМО НА АНГЛИЙСКОМ ЯЗЫКЕ

Фундаментальная и прикладная математика, компьютерные науки

ACADEMIC WRITING IN ENGLISH FOR MATHEMATICS AND COMPUTER SCIENCE

Учебник

Уровни владения языком B2—C1

Р е ц е н з е н т ы :

д-р филол. наук, проф., проф. кафедры английского языкознания филологического факультета МГУ им. М.В. Ломоносова *О.Д. Вишнякова*; канд. филол. наук, доцент, доцент кафедры английского языка как второго переводческого факультета МГЛУ *И.М. Шокина*

А в т о р ы :

доцент департамента иностранных языков
Национального исследовательского университета
«Высшая школа экономики»
М.Б. Антонова;
доцент департамента иностранных языков
Национального исследовательского университета
«Высшая школа экономики»
А.В. Бакулев

Антонова М.Б.

    Учебник нацелен на формирование профессионально ориентированной коммуникативной компетенции в академическом письме в инженерно-технических отраслях науки и устной научной дискуссии. В фокусе учебника — написание проекта исследовательской (курсовой, выпускной квалификационной) работы на английском языке в формате научной статьи для международных рецензируемых журналов, индексируемых в базах данных Web of Science и Scopus. Особое внимание уделяется устной форме научного дискурса — как презентации, так и последующей дискуссии. Материалом учебника послужили статьи, опубликованные в международных рецензируемых журналах.

    Для студентов образовательных программ бакалавриата по направлениям «Информатика и вычислительная техника», «Математика», «Прикладная математика и информатика», «Прикладная математика», «Компьютерная безопасность», «Информационная безопасность», а также для магистрантов и аспирантов соответствующих направлений подготовки.

СОДЕРЖАНИЕ

ПРЕДИСЛОВИЕ

В контексте интернационализации современного научно-образовательного пространства особую роль играет владение иностранным языком, в частности, английским, минимум на уровне независимого пользователя, или Upper-Intermediate (B2) по Шкале Совета Европы. На первый план выходит качественное профессиональное и академическое общение специалистов в условиях международного взаимодействия. Это выражается в участии в конференциях, а также публикации результатов научной деятельности в рецензируемых журналах, индексируемых в базах данных Scopus и Web of Science. В таких условиях особые требования выдвигаются к публикационной активности не только преподавателей, аспирантов и научных сотрудников, но и студентов магистратуры и бакалавриата, что предполагает более высокий минимально необходимый уровень владения английским языком — C1.

С целью соответствия подготовки выпускников вузов вышеуказанным требованиям, в учебных планах отводится место английскому языку для специальных (профессиональных) и академических целей. В ряде университетов предусмотрены написание и защита концепции выпускной квалификационной работы (ВКР) на английском языке. В Национальном исследовательском университете «Высшая школа экономики» (НИУ ВШЭ) концепция ВКР на английском языке (Project Proposal) выполняется в форме научной исследовательской статьи, по формату максимально приближенной к статьям, представляемым к публикации в международных рецензируемых журналах. Подготовка к написанию и защите Project Proposal осуществляется в рамках курса по выбору «Академическое письмо на английском языке», преподаваемого на 4 курсе образовательных программ бакалавриата.

Учебник "Academic Writing for STEM Students" предназначен для студентов инженерно-технических направлений и профилей подготовки, изучающих английский язык для академических целей и призван сформировать релевантные компетенции в области создания письменных академических текстов в формате научных статей, выполненных в логике IMRAD (Introduction — Materials and Methods / Literature Review — Methods — Results and Discussion — (Conclusion)) и защиты научной работы — презентации и участия в научной дискуссии. Ресурс предназначен для студентов, владеющих английским языком не ниже уровня B2. По итогам освоения материалов предполагается повышение уровня владения английским языком до C1.

Учебник разработан для студентов-бакалавров 4 курса факультета компьютерных наук (ФКН) и Московского института электроники и математики им. А.Н. Тихонова (МИЭМ) НИУ ВШЭ, но может использоваться для работы бакалаврами, магистрами и аспирантами других вузов, изучающих математику, информатику и вычислительную технику, прикладную математику, информационную и компьютерную безопасность.

В учебнике важную роль в формировании коммуникативной компетенции в использовании английского языка для академических целей играет коммуникативно-деятельностный подход, в рамках которого широко используется само- и взаимооценивание устно- и письменноречевой продукции студентов, что отвечает современным требованиям гуманизации, индивидуализации и студентоцентричности образовательного процесса.

Учебник состоит из 7 основных разделов (Units), раздела, включающего дополнительные упражнения и материалы (Extra Materials and Activities), ответов к заданиям и упражнений (Keys), справочных материалов и приложений (Appendices) и списка используемых ресурсов (References). В качестве материалов используются фрагменты и полнотекстовые версии оригинальных научных статей, написанных как носителями, так и неносителями английского языка, фрагменты статей на русском языке, а также видеоресурсы и справочные материалы, доступные онлайн. Для удобства студентов ссылки на онлайн-ресурсы представлены как гиперссылками (в случае использования элек-

тронной версии учебника / пособия), так и QR-кодами (в случае использования бумажной версии).

Целью *Unit 1* является знакомство студентов с особенностями научного стиля. *Unit 2* посвящен написанию раздела «Введение» (*Introduction*). *Unit 3* фокусируется на написании обзора литературы (*Literature Review*). В рамках *Unit 4* осуществляется работа над созданием таких разделов, как «Методы» (*Methods*) и «Результаты» (*Results*). *Unit 5* нацелен на написание заключения научной работы, или *Conclusion*. Материалы и задания *Unit 6* сосредоточены на создание аннотации научной статьи (*Abstract*). Цель *Unit 7* — подготовка к устной презентации *Project Proposal* и ответов на вопросы в рамках научной дискуссии. В разделе *Extra Materials and Activities* расположены дополнительные упражнения, отрывки и полнотекстовые версии статей. *Keys* содержит ответы к упражнениям. Блок *Appendices* содержит справочные материалы, критерии оценивания текста и защиты *Project Proposal* и примеры фрагментов и полнотекстовых версий *Project Proposals*, выполненных студентами ФКН и МИЭМ НИУ ВШЭ. В разделе *References* приведен список использованных статей и глав в книгах / сборниках научных трудов, а также ссылки на полезные онлайн-ресурсы.

U n i t 1

FEATURES OF ACADEMIC STYLE

Academic writing is relatively formal. In general, this means that in an essay you should **avoid**:

- vague data
- non-objective statements or overstatements
- colloquial words and expressions: *a lot of, thing, stuff, sort of, to kick the bucket* etc.
- emotional words: *really, very, tremendously* etc.
- abbreviated forms: *can't, doesn't, shouldn't* etc.
- phrasal verbs (two-word verbs): *put off, bring up, get over*
- overusing *There is / There are*-sentences
- overusing the passive voice

Practice 1. The following sentences are mixed formal and informal. Write F (formal) or I (informal) and improve stylistically inappropriate sentences.

1. The project will be completed next year. _____
2. The mob was very rowdy during the protest against cuts to university funding. _____
3. The extracted structures don't contain any simple points and they are topologically equivalent to the original objects. _____
4. It appears that these arguments do not hold water. _____
5. In 1990, Miaoulis, Free and Parsons presented this wonderful technique. _____

6. The technique is perfectly applicable when we have a product that we want to launch on the market. _____
7. If p wouldn't be a border pixel, then a new white component would be arisen by the removal of p, which can't happen because of the simplicity of p. _____
8. Real-world phenomena are very often depicted by graphs where vertices represent entities and edges represent their relationships or interactions. _____
9. The research project won't be continued next year. _____
10. I wonder whether the multiple predictors that influence a response or outcome do so independently or whether they interact. _____

Practice 2. Improve the style of the following sentences.

1. At high level, this architecture relies on the same principles as a a lot of known robotic layered architectures.
2. For more details, we refer to the excellent textbook [74].
3. Not long ago, Goldwasser and Kilian [14] proposed a randomized algorithm based on elliptic curves running in expected polynomial-time, on almost all inputs.
4. Some analysis points out that the volatility in financial time series data is asymmetric.
5. Let's illustrate our proposal on a co-authorship graph depicted in Figure 1.
6. This is a model distribution that postulates that very many innovations are generated by a normal density with a small variance, while too few innovations are generated by a normal density.
7. It has now been well-established that Bayes' rule isn't an apt characterization of how individuals actually respond to new data (Kahneman et a1. [14]).
8. In this survey, we requested average values for full cycle efficiency, durability and some other parameters.

9. The determination of the process noise covariance is generally more difficult as we typically don't have the ability to directly observe the process we are estimating.
10. Generalization of the contradictory information is dangerous and may lead to the creation of a great deal of redundant rules.

Practice 3. Non-objective statements and overstatements may appear, for instance, when introducing research results. It should be borne in mind that any result might face an exception; therefore, it is necessary to hedge statements to show that they do not pretend to cover a hundred percent of all possible cases. Caution can be shown in several ways:

Verbs — **argue, claim, consider, hypothesise, suggest, state, accept, admit, agree, deny, assume, discover, imply, indicate, presume, reveal**

Modal verbs — **may, might, would, can, could**

Adverbs — **apparently, approximately, frequently, generally, hardly, practically, perhaps, probably, possibly, presumably, reasonably, relatively, scarcely, seemingly, slightly, sometimes, typically, usually, virtually**

Adjectives — **probable, possible, certain, uncertain**

IMPERSONAL PHRASES	
It appears to/that	**It is likely that**
It seems to/that	**It is unlikely that**
It would seem to/that	**smth is likely to**
It would appear to/that	**smth is unlikely to**
It is believed that	**It is said that**
It is assumed that	**It has been suggested that**
It tends to	**It is generally agreed that**
There is a tendency to/for	**It is widely accepted that**
Some researchers say that	**It is doubtful if**
Some of the evidence shows that	
It is now generally recognised that	

Find cases of hedging in the following extracts.

A.

The available PPI data are incomplete and often noisy, thus the graphs are generally rather sparse and their edges not very reliable.

(from J.K. Aggarwal et al. *Combinatorial image analysis*)

B.

A third study found norepinephrine response to bicycle ergometry together with psychological factors and blood pressure responses to mental arithmetic to be relatively weak predictors of future blood pressure classification.

(from F. Arnljot et al. *Sympathoadrenal stress reactivity is a predictor of future blood pressure*)

C.

The number of iterations of the loop could be reduced if we could show that a still smaller set of $(X + a)$'s generates a group of the required size. This seems very likely.

...Recently, Hendrik Lenstra and Carl Pomerance [LP2] have given a heuristic argument which suggests that the above conjecture is false. However, some variant of the conjecture may still be true (for example, if we force r > log n).

(from M. Agrawal et al. *Primes in P*)

D.

A product's neighbors are other products that tend to get similar ratings when rated by the same user. For example, consider the movie Saving Private Ryan. Its neighbors might include war movies, Spielberg movies, and Tom Hanks movies, among others. To predict a particular user's rating for Saving Private Ryan, we would look for the movie's nearest neighbors that this user actually rated. As Figure 1 illustrates, the user-oriented approach identifies like-minded users who can complement each other's ratings.

(from Y. Koren et al. *Matrix factorization techniques for recommender systems*)

E.

Figure 2 illustrates this idea for a simplified example in two dimensions. Consider two hypothetical dimensions characterized as female- versus male-oriented and serious versus escapist. The figure shows where several well-known movies and a few fictitious users might fall on these two dimensions. For this model, a user's predicted rating for a movie, relative to the movie's average rating, would equal the dot product of the movie's and user's locations on the graph. For example, we would expect Gus to love Dumb and Dumber, to hate The Color Purple, and to rate Braveheart about average. Note that some movies — for example, Ocean's 11 — and users — for example, Dave — would be characterized as fairly neutral on these two dimensions.

(from Y. Koren et al. *Matrix factorization techniques for recommender systems*)

F.

For instance, a book laying on a furniture might be picked up by Spark and represented in symbolic terms as (BOOK1 type Book, BOOK1 isOn TABLE). These symbolic statements are stored in the knowledge base Oro and made available to the other cognitive modules. Later, the robot might process a sentence like "give me another book". The Dialogs module would then query the knowledge base: find (?obj type Book, ?obj differentFrom BOOK1), and write back assertions like (HUMAN desires GIVE_ACTION45, GIVE_ACTION45 actsOn BOOK2) to Oro. This would in turn trigger the execution controller Shary to prepare to act. It would first call the HATP planner. The planner uses the knowledge base to initialise the planning domain (e.g. find (BOOK2 isAt ?location)), and returns a full symbolic plan to the execution controller. Finally, the controller would execute the plan and monitor its achievement, both for itself and for the human. We present complete examples of similar interactions in Section 4.

(from S. Lemaignan et al. *Artificial cognition for social human — robot interaction: An implementation*)

G.

However, market players are confronted with some questions, for example regarding the limited price difference for the final consumer between peak and valley hours or the lack of experience in the commercial deployment of some of the technologies. This applies even to technologies that seem, theoretically at least, to be cost-effective. In fact, an EU-27 regulatory framework, covering not only power supply, but also energy supply and ancillary services, would be advantageous for the deployment of storage technologies.

(from H.L. Ferreira et al. *Characterisation of electrical energy storage technologies*)

Practice 4. **Make the following sentences cautious using the adverbs *quite*, *fairly* and *rather*. Mind that *rather* helps convey both positive and negative evaluation.**

E.g.: a significant discovery — *quite* a significant discovery
an accurate summary — a *fairly* accurate summary
an inconvenient location — a *rather* inconvenient location
(or: *rather* an inconvenient location)

1. This requires ... substantial cash inflows in each period.
2. The successful match between the ICA-based feature and the traditional category is ... apparent because the circled points are located in the upper part of the diagram.
3. Actually, this is not ... accurate, as we exclude from this value the contribution from all n-grams.
4. The Swedish Sprachbanken project at Göteborgs Universitet offers free access to roughly one billion tokens through a web interface. It is thus ... large.
5. By and large, the orthodox tradition of social science methodology has laid the grounds for a substantial investment in a ... biased view of peace research.
6. The Humean model of the balance of power was first developed for more than two countries and is ... stable under those conditions.

7. It is surprising that ... a large amount of literature is devoted to applications of quantum theory in AI and vice versa, not through quantum computation.
8. Indeed, the relation between quantum automata and automata theory based on quantum logic is ... similar to that between von Neumann's Hilbert space formalism of quantum mechanics and quantum logic.
9. Greedy pursuits, thresholding, and related methods (such as homotopy) can be ... fast, especially in the ultra sparse regime.
10. We found fairly consistent results in average minimum temperature, with about half the sites showing increasing trends in most or all seasons (Table III).

Practice 5. **Academic Style prefers so-called strong verbs to weak ones. Replace the weak verbs in the following sentences.**

Weak verbs: be, give, get, have, make, do
E.g.: He *gave* assistance to my friend (weak verb) — He *assisted* my friend (strong verb).

1. Our approach *makes an attempt* to show that it is possible that much of the linguistic knowledge is emergent in nature and based on specific learning mechanisms.
2. As with the Finnish language, Hungarian *makes use* of a greater number of grammatical cases (usually 18) than German (four cases).
3. He listed some reasons *to give* an explanation why quantum algorithms are so hard to discover.
4. Data mining utilises the algorithm *to do* application, transformation, analysis and further, to identify data's feature and model.
5. Ethnolinguistic fractionalization *has* a direct negative influence on economic growth.

6. This caused a rapid growth of research conducted on the one-species case, whereas the equilibria in two-species case *was* a blank spot.
7. However, even at small *p*, short cuts *are* between distant nodes in the lattice.
8. Our contribution *is* both on the algorithmic and on the complexity side.
9. Most algorithmic papers in sparse reconstruction *give* computational results only on synthetic test problems.
10. In our modified context, there *is* only one implication to consider.

Practice 6. **Try to hedge the following statements.**

E.g.:

a) *I think / believe* that knowledge of Computers and Electronics is a requirement for all computer and mathematical occupations (**wrong**). — *There is little doubt* that / *It seems* that knowledge of Computers and Electronics is a requirement for all computer and mathematical occupations (**correct**).

b) These methods also have potential for describing ecological dynamics in structured landscapes (**wrong**). — The methods also have potential for describing ecological dynamics in *certain kinds of* structured landscapes (**correct**).

1. The probability of three people being friends with each other in a social network, for example, is much higher than would be predicted by a model network constructed using the simple, stochastic process.
2. The design of efficient procedures for searching in a discrete structure is a fundamental problem in discrete mathematics and computer science.
3. The concept lattice, being a universal structure, provides a wealth of information about the relations among objects and attributes.

4. Combining such microscopic and macroscopic properties precisely characterizes the connectivity of the nodes and constitutes information that explains why some vertices have similar attribute values.
5. Depending on the link between vertex attributes and the relationship encoded by the graph, one of these topological properties co-varies with vertex attributes.
6. This empirical evidence demonstrates that thresholding is effective for solving sparse approximation problems in practice; see, e.g., [45]. On the other hand, simulations indicate that simple thresholding techniques behave poorly in the presence of noise.
7. Also, the ellipse error minimization reshapes the ellipses in a bad way, such that they become elongated in space.
8. Figure 7(b) shows the domains, and three stabbings that solve the problem.
9. In either case, whether or not we have a rational basis for choosing the parameters, times superior filter performance (statistically speaking) is obtained by tuning the filter parameters Q and R.
10. Because the training set cannot be considered sparse, looping over each single training case is not practical.

Practice 7. **Rewrite the following passages in a more cautious way.**

A.
Our work opens several perspectives. A short-term perspective is to extend our framework to take into account the information conveyed by categorical vertex descriptors. Another interesting perspective is to adapt the topological pattern mining approach to dynamic graphs by, for instance, identifying unexpected topological patterns over time.

B.
A team of American scientists has found a way to reverse the ageing process. They fed diet supplements, found in health food

shops, to elderly rats, which were then tested for memory and stamina. The animals displayed more active behaviour after taking the supplements, and their memory improved. In addition, their appearance became more youthful and their appetite increased. The researchers say that this experiment is a clear indication of how the problems of old age can be overcome. They state that in a few years' time everyone will be able to look forward to a long and active retirement.

U n i t 2

WHAT MAKES A GOOD INTRODUCTION

Lead-in

Justify the urgency / relevance of the following topics.

1. Credit risk model based on machine learning techniques.
2. Classification of documents based on machine learning.
3. Cloud-based service for training deep neural networks.
4. Automatic classification of factual texts.
5. Sentiment analysis of Twitter data.

Introduction Structure

- Background
  ~ area of research
  ~ relevance of research

- Problem Statement
- Delimitations of the study
  ~ aim / goal
  ~ objectives
- Professional Significance
  ~ theoretical value
  ~ practical value
- Definitions of key terms (if necessary)
- Plan of the paper

Practice 1. Read the Introductions of the articles (see p. 139—145) and indicate the parts of these Introductions according to the plan above.

a) Discovering important bloggers based on analyzing blog threads.
b) Thumbs up? Sentiment classification using machine learning techniques.
c) Recent thinking on economic justice.
d) Quantum computation, quantum theory and AI.

Practice 2. What tenses are used in each part of the Introductions?

Practice 3. Arrange the paragraphs of the Introduction in the correct order.

Mining graph topological patterns: finding co-variations among vertex descriptors

1

Such topological properties and vertex attributes are mostly of numerical or ordinal types and their similarity can be captured by quantifying their co-variation. Such co-variation indicates how a set of vertex descriptors tend to monotonically increase or decrease all together. Therefore, following the way paved by [5], we propose to mine rank-correlated sets over graph descriptors by extracting topological patterns defined as a set of vertex properties and attributes that strongly co-vary over the vertices of the graph.

2

We propose in this article an algorithm, called TopGraph-Miner, which discovers topological patterns. It has as input an attributed graph, such as the one in Figure 1.

3

The same limitation holds for methods [18], [24], [29], [30] that identify sets of vertices that share local attributes and that are close neighbors. Such approaches only focus on a local neighborhood of the vertices and do not consider the connectivity of the vertex in the whole graph. In this paper, we propose to extract meaningful patterns that integrate information about the connectivity of the vertices and their attribute values.

4

We propose several interesting measures of topological patterns that are different from the pairs of vertices considered while evaluating up and down co-variations between properties and attributes: (1) considering all the vertex pairs enables to find patterns that are true all over the graph; (2) taking into account only the vertex pairs that are in a specific order with respect to a selected numerical or ordinal attribute reveals the topological patterns that emerge with respect to this attribute; (3) examining the vertex pairs that are connected in the graph makes it possible to identify patterns that are structurally correlated to the relationship encoded by the graph.

5

Given such a graph, we first compute a set of topological properties for every of the graph vertices. TopGraphMiner then integrates searching and pruning strategies in the identification of the most relevant topological patterns.

6

Real-world phenomena are often depicted by graphs where vertices represent entities and edges represent their relationships or interactions. Entities are also described by one or more attributes that constitute the attribute vectors associated with the vertices of the attributed graph.

7

We also propose an operator that identifies the top k representative vertices of a topological pattern. Let us illustrate our proposal on a co-authorship graph depicted in Figure 1, where vertices (from A to P) denote authors, edges encode co-authorship relations, and three attributes describe the author.

8

This article is structured as follows: Section 2 presents topological vertex properties. Sections 3 and 4 introduce our new model for mining topological patterns. Our algorithm is defined in Section 5. Its efficiency and its effectiveness are shown in Sections 6 and 7. Section 8 discusses the related work and Section 9 concludes the article.

9

Our contribution is therefore threefold. First, we propose a new kind of graph analysis that exploits attributes and topological properties of vertices.

10

Existing methods that support the discovery of local patterns in graphs mainly focus on the topological structure of the patterns, by extracting specific subgraphs while ignoring the vertex properties (cliques [22], quasi-cliques [21], [32]), or compute frequent relationships between vertex attribute values (frequent subgraphs in a collection of graphs [16] or in a single graph [4]), while ignoring the topological status of the vertices within the whole graph, e.g. the vertex connectivity or centrality.

11

Second, to produce such analysis, we provide new insights into co-variation pattern mining by considering up and down co-variations,

defining new upper bounds on the support of such co-variations, proposing several interestingness measures of topological pattern, and giving to the user the ability to visualize the patterns on the original graph thanks to the identification of the top k representative vertices. Third, to validate our approach, we conducted an empirical study.

12

Combining such microscopic and macroscopic properties precisely characterizes the connectivity of the nodes and constitutes information that may explain why some vertices have similar attribute values. Depending on the link between vertex attributes and the relationship encoded by the graph, one of these topological properties may co-vary with vertex attributes.

13

Finally, it gives to the user the ability to visualize every pattern on the input graph by identifying the top k representative vertices.

PUNCTUATION FOCUS

Practice 4. Pick the terms for the punctuation marks in the chart.

ellipsis	comma	semi-colon	question mark	hyphen colon
dash	square brackets	full stop / period		parentheses / brackets
apostrophe	quotation marks / inverted commas			exclamation mark / point
braces	(forward) slash	at sign / symbol	underscore	asterisk
ampersand	hash(tag) / number sign	back slash		

@	#
( )	{ }
' (John's room)	/
!	?
…	.
— (singer-songwriter)	*
;	:
_ (Paper_Smith)	— (— Where were you?)
&	[ ]
'	" … "
\	

Practice 5. Go to Appendix B Study the main rules of English punctuation; then punctuate the following sentences.

A.

1. In UT2004 a human player uses a keyboard and a mouse to control the actions of an avatar in a 3-D confined virtual world.

2. Therefore in order to make our agents generalized and at the same time to ensure they learn appropriately and effectively we propose to combine the fuzzy ART and ART2 operations.

3. Moreover no matter how β is set a code will become redundant eventually if contradictory input vectors keep coming in. Because once a code is over-generalized it will always fail the vigilance criterion defined in Eq. (3).

4. If this can be established the alternative view emerges that dictionary knowledge is a subset of more general encyclopaedic knowledge.

5. According to this view it is only the dictionary component that properly constitutes the study of lexical semantics the branch of semantics concerned with the study of word meaning.

6. In cognitive semantics the distinction between lexical and compositional semantics is not seen as a useful division.

7. This was important work but it did contain one flaw: no distinction was made between information about the world that stems from perception, action or the internal feelings involved in actions.

8. The latter point might seem to negate the first analog spatial representations may be built from all kinds of perceptual information but they cannot exist in all modalities and be spatial abstractions at the same time. In any case this was the view that was widely accepted by cognitive linguists.

9. The location of the boundary of a category is independent of its prototype that is to say two categories may have the same prototype but different boundaries likewise two categories may have the same boundaries but different prototypes.

10. In Theoretical Linguistics researchers sometimes try to obviate limitations of available resources through Googleology. When data on low-frequency or non-standard phenomena is needed search engine queries (mostly using Google's service) are used to look for single occurrences of some grammatical construction, or even worse result counts returned for such queries are used for more or less formal statistical inference.

11. Such tension also exists in other economic contexts most prominently in the public finance literature.

12. In this section we explore the thesis that semantic structure encodes and externalises conceptual structure.

13. If, on the other hand we choose the combination A+C we are increasing the overall product penetration by 1 individual because C has been chosen by a consumer that did not choose A. Our overall penetration will be 4 the maximum that we could reach.

14. Therefore following the way paved by [5] we propose to mine rank-correlated sets over graph descriptors by extracting topological patterns.

15. In Section 4 we state the algorithm and present its proof of correctness.

16. A series of simple regressions cannot accomplish these tasks if we want to examine the simultaneous effects of multiple predictors on a response we must use a method that treats them accordingly.

17. In 1960 R.E. Kalman published his famous paper describing a recursive solution to the discrete-data linear filtering problem [Kalman60]. Since that time due in large part to advances in digital computing the Kalman filter has been the subject of extensive research and application particularly in the area of autonomous or assisted navigation.

18. In this paper we look in detail at three particular networks of scientific collaborations and describe some of the patterns they reveal.

19. In the context of the present example this means in part that there is no relationship between quantitative and verbal ability. In more advanced models of factor analysis the condition that the factors are independent of one another can be relaxed. As for the factor means and variances the assumption is that the factors are standardized.

20. To overcome these drawbacks the construction of storage tanks has been proposed.

B.

1. In this work we carry out an extensive empirical comparison of the most widely-used smoothing techniques.

2. Due to this phenomenon Mandelbrot argues that such variance might change over time. In response Engle [9] proposed the Autoregression Conditional Heteroskedasticity (ARCH) model to solve for the biases of traditional econometrics models.

3. In Eq. 4 $r(t)$ is the threshold value used in this model as this value is usually complex and cannot be solved for analytically we therefore use a fuzzy model to solve it.

4. The basic configuration of fuzzy systems consists of four components a fuzzifier a fuzzy rule base a fuzzy inference engine and a defuzzifier.

5. In the remainder of this section we discuss some knowledge-intensive methods.

6. With the growth of numbers of users and items the system needs more resources for processing information and forming recommendations.

7. This in turn would require development of more specified recommendation systems.

8. Typically quantum computation is realized by quantum circuits consisting of quantum gates.

9. Research arising from the interplay between quantum theory and AI can be roughly classified into two categories (1) using some ideas from quantum theory to solve certain problems in AI and (2) conversely applying some ideas developed in AI to quantum theory.

10. As will be shown later it is straightforward to sign the direction of any inference bias.

11. A primary objective of this paper therefore is to prevent these type 1 errors from occurring in the future.

12. Specifically it is likely that some valid research is not making it through the publication process because authors or referees are incorrectly focusing on statistically insignificant interaction term coefficients rather than on potentially significant differences in marginal effects. A primary objective of this paper therefore is to prevent these type 1 errors from occurring in the future.

13. In the first set of three trials we utilize a turnover function.
14. For this purpose three types of indices are used: index of particular words, indices of parts of speech, and index of patterns yet matched.
15. In [7] for problems with the information in the form of messages about the parametric importance of criteria a sufficient condition that turns inclusion (6.1) into equality was given this condition assumes that there are continuous criteria with closed intervals bounded from both sides.
16. For the sake of simplicity we assume that all the discrete criteria in the problem are integer that is we assume that the values in their scales are sequences of integers.
17. Then for the initial vector v one can go along the tree branches from the leaves to the root while deteriorating and improving the corresponding tree estimates.
18. To repeat our goal is to test whether the overreaction hypothesis is predictive.
19. In summation neural network and logistic regression methodology outperform other models.
20. When taking the gold standard splits from our reference set of 700 words our program gathered 216 periphrase candidates from the corpus of which 160 were correct a precision of 74%.

GRAMMAR FOCUS

Practice 6. **Insert the appropriate article where necessary and punctuate the sentences if necessary.**

1. In ... Section 2 of this article we describe our main test data set, ... corpus in five languages. In ... Section 3 we describe ... scheme based on readers' judgements that enables ... setting up of a similarity space.
2. In ... most situations neither Eqs. (7) nor (8) are equal to b_3.
3. In ... context of ... turnover literature focusing solely on ... estimated b_3 coefficient could lead to ... inference that ... marginal effect of performance differs with ... type of ... firm.
4. ... parameter b is what generates ... strength of ... relationship between turnover and performance.
5. ... solution, also demonstrated by ... way of simulations, is to separately calculate ... marginal effect of ... change in performance for the two types of firms.
6. These are ... elements that made ... neural network.
7. ... greater ... link value, ... easier it is for the link to become too activated and affect ... neural network.
8. ... Figure 3 shows ... working process of this study.
9. ... Australia credit database has 690 case data.
10. Thanks to ... rapid development of information technology, researchers and financial institutions began to use ... data mining and machine learning.
11. ... word can belong to several syntactic categories simultaneously.
1. When we crossed ... threshold of connecting more objects than people to ... Internet, ... huge window of opportunity opened for ... creation of applications in ... areas of automation, sensing, and machine-to-machine communication.
12. For each group of patterns, ... automatic recognition procedure was developed and experimentally studied.
13. ... directions for ... future work include ... out-of-core implementations of ... presented simplification methods.
14. Nevertheless there is one phenomenon that seems to adversely and significantly affect ... performance of ... certain group of algorithms on ... small number of datasets.

Practice 7. **Translate the sentences with the phrase *Given (that)*, which introduces a known fact.**

1. Given living trees absorb CO2 (carbon dioxide) and produce oxygen, reducing the tree population further compounds the problem.

2. Given his age, he's remarkably active.
3. Given the hypothesis that semantic structure reflects conceptual structure, the system of semantic structure is also divided into two subsystems.
4. Given the diverse applications of web corpora, users have to make the design decisions themselves, based on our description of how the procedures work.
5. Given high interannual and decadal variability in climate, it is obvious that short-term climate records can be misleading in evaluating trends.
6. Given the current state *s*, TD-FALCON first decides between exploration and exploitation by following an action selection policy.
7. This allows us to ask the probability of two scientists coauthoring a paper, given that they have a third mutual collaborator and have not collaborated in the past.
8. Given such graph, it first computes a set of topological properties for every of its vertices.
9. It was surprising the government was re-elected, given that they had raised taxes so much.
10. Given a set of objects of a domain, attribute exploration builds an implicational theory of the entire domain.

Practice 8.

A. Read the Introductions of the following articles (see p. 139—143): *Discovering important bloggers based on analyzing blog threads; Thumbs up? Sentiment classification using machine learning techniques; Recent thinking on economic justice*.

Find words / phrases that you would like to use when writing introductions of your project proposals. Then write these words / phrases in the following chart.

Discovering important bloggers based on analyzing blog threads	Thumbs up? Sentiment classification using machine learning technique	Recent thinking on economic justice
1. ubiquitous 2. to treat smth as 3. ...	1. a challenging aspect 2. apart from 3. ...	1. In light of space limitation 2. to impose constraints 3. ...

B. Go to p. 180—181 and compare your list of words / phrases with the list given. Explain the meaning of these words / phrases or give their synonyms.

Practice 9. **Read the extracts from the article by E. Pazner *Recent thinking on economic justice* and find the following words / phrases, explain their meaning or give their synonyms. Then make up sentences with these words / phrases so that to include them in the introduction to your project proposals.**

- to avoid terminological ambiguities
- to impose constraints
- smth is said to be...
- a well-established result (is that...)
- smth is referred to as // smb refers to smth as...
- this result holds only under the assumption that...
- On the understanding that...
- On the assumption that...
- Under the assumption on smth
- The 1st question that arises is whether smth is/are...
- question of interest to smb
- At this point it is important to...

- This follows from the fact that...
- The situation drastically changes when...
- This implies that...
- Insofar as smth is concerned...
- Virtually
- The polar extreme to this view that... is... // The polar extreme to the view is...
- in terms of // regarding
- the above-mentioned smth
- likewise
- to shed light on sth
- It is worth emphasizing that...
- in light of the known difficulties
- To sharpen the issue
- smth is sought
- a similar remark applies to...
- The main contribution of...
- smth ought to be sought

A.

To avoid terminological ambiguities, let me briefly review some standard economic terms.

The term allocation means a list of consumption bundles, one for each individual, satisfying the feasibility (scarcity) constraints imposed by the aggregate consumption possibilities. In the case of production economies, it is worth reminding that an allocation includes a full description of the consumption of leisure by each individual in addition to his consumption of "ordinary" consumer goods.

An allocation is said to be Pareto-efficient if it is impossible to make one individual better off by reshuffling consumption and/or production activities among individuals without making at least one other individual worse off.

A well-established result in welfare economics is that given any initial allocation there exist prices such that trading at these prices leads to a Pareto-efficient allocation at which each market is cleared. The initial allocation is usually referred to as the initial endowments (i.e. a list of initial holdings of commodities by each consumer.) This result holds only under a set of standard assumptions on the nature of the economic environment (i.e. preferences and technology). On the understanding that each trader makes a utility maximizing trading decision (subject to his budget or income constraint defined by the value of his initial endowment at the prevailing prices), and on the assumption that each is unable to influence the prevailing prices (which are the same as those faced by every other trader), the underlying market structure is said to be perfectly competitive. The market clearing allocation is called a competitive equilibrium. The next two sections are a review of recent attempts at defining allocations that are both just and efficient. Such direct attempts at assessing the equity of allocations in terms of their intrinsic properties is what distinguishes the theories reviewed here from the traditional approaches to distributive justice.

B.

The first question that arises is whether envy-free allocations are always possible. Noting that any egalitarian allocation (in which every consumer is assigned an identical consumption bundle) is obviously envy-free, the answer to this first question is affirmative. The next question of interest to any economist raised in the tradition of modem Paretian welfare economics relates to whether among all fair allocations at least one can be found that is also Pareto-efficient. On the assumption that Pareto-efficiency is a normatively appealing desideratum, an affirmative answer to this second question would establish the possibility of satisfying the equity requirement of fairness without having to give up the separate objective of economic efficiency.

At this point it is important to distinguish between exchange and production economies. This can easily be seen by noting that if each consumer is assigned an identical initial endowment of commodities and is allowed to trade in perfectly competitive markets, the resulting

competitive equilibrium allocation must be envy-free. This follows from the fact that at such an allocation every consumer faces the same prices, has the same income and competitive behavior implies that each consumer's utility is maximized subject to his budget constraint.

The situation, however, is drastically changed when production is allowed and when it is realistically assumed that individuals are not identically endowed with innate productivity skills so that different individuals contribute differently to social product per unit of time input of their labor skills. This implies that the argument used above for exchange. Economies cannot be advanced in the case of production economies. And indeed it is now known that production economies satisfying the standard assumptions of economic theory can be constructed in such a manner that among all Pareto-efficient allocations none can be found that is envy-free. This is a disturbing result insofar as the equity notion of fairness is concerned since economic efficiency is the normative criterion upon which the foundations of modern welfare economics are laid. Virtually every economist will agree that the normative appeal of any equity notion that stands in conflict with Pareto-efficiency should be doubted.

C.

The polar extreme to the view that one has an inherent right to the fruits of one's own labor is the position that individual productivity skills belong to no one in particular and ought to be regarded as a common pool of valuable resources which belongs to society as a whole.

In terms of the coordination of economic activity by means of a market-like mechanism capable of achieving potentially-fair and efficient allocations, a good way to think about this is the following. Imagine that the state "collects" everybody's initial endowment of time and every other available resource in the society. In compensation, each individual gets the same amount of abstract purchasing power by means of which he can buy resources from the state at competitive prices. One problem with both of the above-mentioned (modified) fairness criteria is that each systematically discriminates among individuals in line with their respective productive abilities.

Consider the set of possible allocations, i.e. the set of all lists of consumption-cum-leisure bundles that the economy under consideration can provide. Out of the set of all possible allocations the efficiency criterion chooses those that are Pareto-efficient. Thus Pareto-efficiency criterion induces a restriction on the set of allocations we are willing to admit from a normative point of view. Likewise, any equity requirement restricts the set of normatively admissible allocations to those satisfying the equity requirement in question.

D.

I wish to conclude with some brief remarks on the way the approaches reported here might shed light on some major issues of social policy. It is worth emphasizing that any notion of economic justice must be based on, or start from, an underlying notion of individual rights in the social sphere. In the case of the original fairness criterion and the equivalence criteria discussed above the starting point is that "rational" individuals have an inviolable right to their preferences. The problem of social justice is then viewed as searching for an equitable way to "aggregate" individual preferences to overcome the fundamental conflict of interest arising from the fact that economic resources are scarce relative to individuals' wants. In this sense, the approaches discussed here are consistent with the mainstream of modern welfare economics and its underlying sanction of individual preferences.

Their contribution ought to be judged in light of the known difficulties involved in the search for a "good" aggregation procedure. I will attempt now to elucidate briefly the precise sense in which individuals are considered here to have a right to their preferences and to illustrate thereby how the present approach might be translated into concrete social policy recommendations. To sharpen the issues, suppose that all individuals have exactly

identical preferences. There is absolutely no need for a tax-transfer mechanism in this instance on equity grounds if the original notion of fairness is accepted. If on the other hand satisfaction of the egalitarian equivalence criterion is sought, a tax-transfer scheme will be required even in this case.

The more difficult case is that where individuals differ in both preferences and abilities. In this case tax-subsidy schemes will generally be required even for the attainment of Pareto-efficient and envy-free allocations (whenever such exist). A similar remark applies to satisfaction of either of the equivalence criteria. Note also that the notion of equality of opportunity is more difficult to define in this general case as there is no solid yardstick of equality to start with since individuals are different in every relevant dimension. These difficult issues are the subject of current research.

I conclude by noting that it is my belief that the main contribution of the recent developments reported here might ultimately prove to lie in their enabling the discussion of policy issues within a well-defined conceptual framework. Legitimate differences in opinions will then properly reflect differences in underlying values as opposed to methodological disagreement on the proper way to approach the policy dilemma in the first place. It is on. And only on, this latter aspect that professional agreement ought to be sought.

Practice 10. **Read the extracts from the articles and find the following words / phrases, explain their meaning or give their synonyms. Then make up sentences with these words / phrases so that to include them in the introduction to your project proposals.**

- It is also assumed that
- be subject to smth
- In this framework
- Given
- to get some idea about how ...
- provisional conclusion
- insight

- tentative conclusion / estimate
- crude estimate
- be pioneered by
- an umbrella term for
- feasible
- account for

A.

Bidder i knows the realization x_i of X_i and only that other bidders' values are independently distributed according to F. Bidders are risk neutral— they seek to maximize their expected profits. All components of the model other than the realized values are assumed to be commonly known to all bidders. In particular, the distribution F is common knowledge, as is the number of bidders.

Finally, it is also assumed that bidders are not subject to any liquidity or budget constraints — each bidder i has sufficient resources so that, if necessary, he or she can pay the seller up to his or her value x_i. Thus, each bidder is both willing and able to pay up to his or her value.

We emphasize that the distribution of values is the same for all bidders and we will refer to this situation as one involving symmetric bidders. In this framework, we will examine two major auction formats:

a. A first-price sealed-bid auction, where the highest bidder gets the object and pays the amount he bid.
b. A second-price sealed-bid auction, where the highest bidder gets the object and pays the second highest bid.

Each of these auction formats determines a game among the bidders. Given that bidders are symmetric, it is natural to focus attention on symmetric equilibria.

In a first-price auction, equilibrium behavior is more complicated than in a second-price auction. Clearly, no bidder would bid an amount equal to his or her value since this would only guarantee a payoff of 0. Fixing the bidding behavior of others, at any bid that will neither win for sure nor lose for sure, the bidder faces a simple

trade-off. An increase in the bid will increase the probability of winning while, at the same time reducing the gains from winning. To get some idea about how these effects balance off, we begin with a heuristic derivation of symmetric equilibrium strategies.

(from V. Krishna *Auction Theory*)

B.

In this paper we have modelled a mathematics MOOC forum discussion as a network and we have resorted to network analysis and to engagement structures to analyse the data. The aim of this study, with a limited number of cases, is not to draw general conclusions about the relationships between each participant's desire in the discussion and the kind of network that results from them. Instead, we aim at discussing within the MAVI community the viability of applying such lenses of analysis to a discussion forum and show possible, provisional conclusions about the kind of insight we can get from this.

(from B. Rott et al. *Views & beliefs in mathematics education: The role of beliefs in the classroom*)

C.

Thus we can tentatively conclude that the internal fields for a uniform anyon fluid in the AFM case are undetectable by μSR. We have used our firm results for $M_0(T)$ as a basis for some other, very tentative estimates. A very crude estimate of the field from a local moment, which is induced by the charge inhomogeneities which screen the charged muon probe, gives a value exceeding the experimental bound by approximately an order of magnitude.

(from C.S. Ting et al. *Physics and Mathematics of Anyons*)

D.

Game theory is the science of strategy. It attempts to determine mathematically and logically the actions that "players" should take to secure the best outcomes for themselves in a wide array of "games." Game theory was pioneered by Princeton mathematician John Von

Neumann. In the early years the emphasis was on games of pure conflict (zero-sum games). Today, game theory applies to a wide range of behavioral relations, and is now an umbrella term for the science of logical decision making in humans, animals, and computers.

(from A. Dixit & B. Nalebuff *Game Theory*)

E.

The set of feasible designs is sometimes referred to as the feasible region, especially for optimization problems with two design variables. It is important to note that the feasible region usually shrinks when more constraints are added to the design model and expands when some constraints are deleted.

(from J.M.F. Ten Berg & G. Sočan *The set of feasible solutions for reliability and factor analysis*)

F.

Spam accounts for 14.5 billion messages globally per day. In other words, spam makes up 45% of all emails. Some research companies estimate that spam email makes up an even greater portion of global emails, some 73% in fact. The United States is the number one generator of spam email, with Korea clocking in as the second largest contributor of unwanted email.

(from *Spam Laws*)

G.

This paper examines the naïve Bayesian model and extensions of it to account for the effects of inverse base rates. These are human categorization phenomena in which base rate information appears to be ignored. The naïve Bayesian classifier accounts for a subset of the phenomena observed in base rate experiments. An extension to the model is examined that uses structure in tha data sets resulting from features shared between categories.

(from L.J. Frey et al. *Naïve Bayesian accounts of base rate effects in human categorization*)

Practice 11. Complete the extracts with the words / phrases from the boxes. Highlight words / phrases that account for cautiousness.

A.

| precise   would   briefly   conversely   to flesh this out a bit   into account |
| encompassing   introduction   refer   the work still to be done |
| so as   in words |

Advanced versions of attribute exploration can take 1) .... background information, such as relations among attribute values (not necessarily given as implications), thus, avoiding suggesting trivial implications [5]. There are also methods of concept exploration, relational exploration, and rule exploration that, each in its own way, generalize attribute exploration, making it possible to work with a broader class of dependencies [10], [11], [12]. We 2) .... to all such methods collectively as *conceptual exploration*.

Our aim is to extend these methods 3) .... to make them suitable for collaborative ontology construction over the web by a geographically spread community of researchers working in the same domain. The idea is to provide tools that 4) .... allow them to use attribute exploration and related techniques to refine the language in which their domain is described and boost their knowledge about it. Being suggested, an implication will be accepted only if no expert has a counterexample for it. In the end, there will be a list of implications correctly describing objects under study and a representative context.

We start with a short 5) .... into the relevant aspects of formal concept analysis and then define attribute exploration. After that, we discuss possible ways to make this procedure collaborative. Finally, we describe the current state of our web-based exploration system and 6) .... We 7) .... introduce necessary mathematical definitions and then explain them less formally.

Given a (formal) context K = (G; M; I), where G is called a set of objects, M is called a set of attributes, and the binary relation I ⊆ G × M specifies which objects have which attributes, the derivation operators $(\cdot)^I$ are defined for I ⊆ G and B ⊆ M as follows:

$$A^I = \{m \in M \mid \forall\ g \in A : gIm\}$$
$$B^I = \{g \in G \mid \forall\ m \in B : gIm\}$$

8) ...., A^I is the set of attributes common to all objects of A and B^I is the set of objects sharing all attributes of B.

The formal context makes 9) .... the scope of the discussion by specifying the domain to which it applies (listing all the objects of this domain) and defining the terms in which it is going to be discussed (listing the attributes to be used in object descriptions). 10) ...., we give a small example based on the data from the O*NET Resource Center, which essentially provides an interface to a taxonomy of occupations, organizing occupations in various groups and describing the knowledge, skills, and abilities required by each occupation. Here, we focus on the knowledge required by occupations from the Computer and Mathematical job family. A formal context 11) .... three occupations is shown in Fig. 1.

A line diagram of the concept lattice of this context is shown in Fig. 2. Nodes correspond to formal concepts, with more general concepts placed above less general ones. Two concepts are connected with a line if one is more general than the other and there is no concept between the two. Every concept in the diagram is described extensionally, by a group of objects, and intensionally, by attributes shared by all the objects in the extent of this node. The names of the objects in the extent of a node can be read off from the diagram by looking at the labels immediately below this node and below all nodes that can be reached from this node by downward arcs. 12) ...., the set of attributes forming the intent of a node consists of labels immediately above this node.

(from S. Obiedkov et al. *Collaborative conceptual exploration as a tool for crowdsourcing domain ontologies*)

B.

introduce threefold to validate discusses which exploits conducted presents associated new insights given wherein encode

Let us illustrate our proposal on a co-authorship graph depicted in Figure 1, 1) .... vertices (from A to P) denote authors, edges 2) .... co-authorship relations, and three attributes describe author: h corresponds to the author h-index, which attempts to measure both the productivity and the impact of the published work of each author [15]; i denotes the average number of hours per week spent by each author on instructional duties; and t designates the number of publications the author had in the IEEE TKDE journal. As topological property, we consider the betweenness centrality measure that is the number of times a vertex appears on a shortest path of the graph (see Section 2). This value is in a circle 3) .... to each vertex on Figure 1.

We propose in this article an algorithm, called TopGraph-Miner, 4) .... discovers topological patterns. It has as input an attributed graph, such as the one in Figure 1. 5) .... such graph, it first computes a set of topological properties for every of its vertices. TopGraphMiner then integrates searching and pruning strategies in the identification of the most relevant topological patterns. Finally, it gives to the user the ability to visualize every pattern on the input graph by identifying the top k representative vertices.

Our contribution is therefore 6) .... First, we propose a new kind of graph analysis that 7) .... attributes and topological properties of vertices. Second, to produce such analysis, we provide 8) .... into co-variation pattern mining by considering up and down co-variations, defining new upper bounds on the support of such co-variations, proposing several interestingness measures of topological pattern, and giving to the user the ability to visualize the patterns on the original graph thanks to the identification of the top k representative vertices. Third, 9) .... our approach, we 10) .... an empirical study that includes: (1) a comparison of TopGraphMiner with a baseline approach; (2) a study of its

empirical complexity; (3) an analysis of the pruning capability of the proposed upper-bound; (4) we provide results on the execution time with and without the pruning strategy; and (5) we study qualitatively some patterns extracted from four real-life networks: a co-authorship network, a movie cofactor network, a patent citation network, and gene interaction network.

This article is structured as follows: Section 2 11) .... topological vertex properties. Sections 3 and 4 12) .... our new model for mining topological patterns. Our algorithm is defined in Section 5. Its efficiency and its effectiveness are shown in Sections 6 and 7. Section 8 13) .... the related work and Section 9 concludes the article.

(from A. Prado et al. *Mining graph topological patterns: finding co-variations among vertex descriptors*)

C.

distortions poses compare involved result heuristics similar substantial converted corrupted key capable constitute

Irregularly sampled point clouds 1) .... one of the canonical input data formats for scientific visualization. Very often such data sets 2) .... from measurements of some physical process and are 3) .... by noise and various other 4) .... Point clouds can explicitly represent surfaces, e.g. in geoscience [12], volumetric or iso-surface data, as in medical applications [8], or higher dimensional tensor fields, as in flow visualization [22]. For surface data acquisition, modern 3D scanning devices are 5) .... of producing point sets that contain millions of sample points [18].

Reducing the complexity of such data sets is one of the 6) .... preprocessing techniques for subsequent visualization algorithms. In our work, we present, 7) .... and analyze algorithms for the simplification of point-sampled geometry.

Acquisition devices typically produce a discrete point cloud that describes the boundary surface of a scanned 3D object. This sample set is often 8) .... into a continuous surface representation, such as polygonal meshes or splines, for further processing. Many of these

conversion algorithms are computationally quite 9) .... [2] and require 10) .... amounts of main memory. This 11) .... great challenges for increasing data sizes, since most methods do not scale well with model size. We argue that effective surface simplification can be performed directly on the point cloud, 12) .... to other point-based processing and visualization applications [21, 1, 14].

Different 13) .... have been presented in the polygonal mesh setting (see [9] for an overview) that we have adapted and generalized to point-based surfaces.

<div align="right">(from M. Pauly et al. Efficient Simplification
of Point-Sampled Surfaces)</div>

D.

afterwards   heuristically   runs   proposed   required   ultimate developed   so far   independently   fundamental   achieved   based property   to yield   let

Prime numbers are of 1) .... importance in mathematics in general, and number theory in particular. Of special interest are those properties that allow one to determine efficiently if a number is prime. Such efficient tests are also useful in practice: a number of cryptographic protocols need large prime numbers.

2) .... PRIMES denote the set of all prime numbers.

In 1975, Miller [Mil] used a 3) .... based on Fermat's Little Theorem to obtain a deterministic polynomial-time algorithm for primality testing assuming the Extended Riemann Hypothesis (ERH). Soon 4) .... his test was modified by Rabin [Rab] 5) .... an unconditional but randomized polynomialtime algorithm. 6) ...., Solovay and Strassen [SS] obtained, in 1974, a different randomized polynomial-time algorithm. Since then, a number of randomized polynomial-time algorithms have been 7) .... for primality testing, based on many different properties.

In 1983, Adleman, Pomerance, and Rumely 8) .... a major breakthrough by giving a deterministic algorithm for primality that runs in $(\log n)^{O(\log \log \log n)}$ time (all the previous deterministic

algorithms 9) .... exponential time). Their algorithm was (in a sense) a generalization of Miller's idea and used higher reciprocity laws. 10) .... on their ideas, a similar algorithm was 11) .... by Atkin [Atk]. Adleman and Huang [AH] modified the Goldwasser-Kilian algorithm to obtain a randomized algorithm that 12) .... in expected polynomial-time on all inputs.

The 13) .... goal of this line of research has been, of course, to obtain an unconditional deterministic polynomial-time algorithm for primality testing. Despite the impressive progress made 14) ...., this goal has remained elusive. In this paper, we achieve this. We give a deterministic, $\tilde{O}(\log^{15/2} n)$ time algorithm for testing if a number is prime. 15) ...., our algorithm does better.

<div align="right">(from M. Agrawal et al. Primes in P)</div>

Practice 12. Render the following Russian sentences in English using the vocabulary discussed in this Unit.

A.

1. Предлагаемый метод ***позволяет*** решить эту сложную задачу.
2. ***Помимо того, что*** фермеры столкнулись с серьезными организационными трудностями, они вынуждены платить довольно-таки высокие налоги.
3. Применение этого алгоритма зависит от ***целого ряда*** факторов.
4. Присутствовало ***много*** людей.
5. ***Количество*** участников было больше летом.
6. В этой компании отличные возможности ***в плане*** карьерного роста.
7. Данный результат ***верен*** только при определенных условиях.
8. Чтобы избежать терминологических ***разночтений***, автор ***кратко напоминает*** некоторые термины.
9. В своей статье автор ***ограничивается*** только этим вопросом.

10. Поскольку эта проблема *многоаспектная*, автор предлагает одно из возможных решений.
11. *Поскольку мы ограничены рамками настоящей статьи*, приведем лишь *результаты*.
12. Были предприняты *попытки* найти другое решение данной проблемы.
13. *Первый вопрос, который возникает*, это *является ли* данное решение лучшим по сравнению с предыдущими.
14. Эксплицитно выраженный фидбэк пользователей *считается* рейтингами.
15. Данные ограничения, по-видимому, являются *повсеместными* для этой группы уравнений.

B.

1. *Прямо противоположным* традиционному является новый подход, предлагаемый в этой статье.
2. Наши эксперименты *пролили свет на* то, как функционирует данное правило в этом множестве.
3. Автор *называет* эти методы концептуальной классификацией.
4. Этот результат не *подвержен* ограничениям.
5. Для *подтверждения* своего подхода, мы использовали следующие теоремы.
6. Результаты исследования показывают, что признаки A превалируют. *Напротив*, количество признаков B наименьшее.
7. *Чтобы получить некоторое представление о* том, как работает этот алгоритм, рассмотрим следующий пример.
8. Помимо того, что они обозначили это явление, они *предложили новый взгляд на* него.
9. *Предстоящая работа* описывается в Заключении.
10. В данной статье ставятся *две* цели.
11. *Учитывая* эти отклонения, возможно вывести предполагаемую погрешность.

12. Авторы *рассматривают* этот вопрос как наиболее важный.
13. *То, что касается* этой модели, ее возможно применить в нескольких случаях.
14. Компаниям приходится отвечать разнообразным требованиям клиентов, *достаточно назвать лишь некоторые из них*: ясность контрактов, краткие сроки, невысокая стоимость продукта.
15. Способы построения этой модели *еще предстоит найти*.

Practice 13.

A. Read the following Active Voice English sentences and translate them into Russian. Pay attention to the differences in syntactic structure of the corresponding English-Russian sentences.

E.g.:
   a) **Section 2 reviews** the recent economic theory. — **В разделе 2** приводится обзор современной экономической теории.
   b) **2018 witnessed** a skyrocketing of individual investments in securities. — **В 2018 году** произошел существенный рост объемов инвестиций населения в ценные бумаги.

1. Section 4 presents some clarifying comments.
2. This section briefly surveys previous work on non-topic-based text categorization.
3. A short Bibliographical Addendum concludes the article.
4. Figure 2 illustrates this idea for a simplified example in two dimensions.
5. As Figure 1 illustrates, the user-oriented approach identifies like-minded users who can complement each other's ratings.
6. Recent years have seen rapid growth in on-line discussion groups and review sites.
7. The crash killed 106 people.
8. The advertisement says the rent is 700 pounds a month.

9. Russian history had not seen serious iconoclastic movements until the short period after the revolution of 1917.
10. Therefore, these models treat the item bias b_i as a function of time.

B. Render the following Russian sentences in English applying the above syntactic patterns.

1. За прошедшие 10 лет в городе значительно увеличился процент преступности.
2. В прошлом году в столице произошло несколько массовых выступлений.
3. В последние годы наблюдается сильный рост числа малых фирм.
4. За последние годы произошел резкий рост инфляции.
5. В этом методе учитываются не все возможные факторы, влияющие на конечный результат.

Practice 14. Render the following Introductions from Russian into English.

A.

Ландшафт современного образования формируется в том числе и под воздействием информационных технологий, предлагающих разнообразные способы доставки учебного материала, новые формы общения и взаимодействия участников образовательных процессов, а также — во многих случаях — существенно снижающих издержки на поддержку этих процессов [1].

Общим термином для обозначения учебного процесса, часть которого организована при помощи компьютерных информационных систем (информационных систем, ИС) стало гибридное обучение, содержательное определение которого можно найти, например, в [2]. Наиболее интересными представителями ИС за последние пять лет стали развившиеся до полноценных платформенных решений информационные системы, позволяющие поддерживать процессы гибридного обучения, особенно требующие организации и поддержки взаимодействия многих участников при работе в общем пространстве над общими артефактами.

В статье мы коснемся коллаборативных ИС поддержки гибридного обучения. Очевидно, что появление их последнего поколения, затронуло не только процесс обучения, но и способы управления им. Статья сфокусирована на технологических возможностях, позволяющих как преподавателям, так и администраторам трансформировать свои рабочие процессы и повысить их эффективность.

(from O. Maksimenkova et al. *Collaborative Technologies in Education: How to Build an Effective System for Blended Learning?*)

B.

Многоагентный подход широко используется при создании программных средств информационных систем, отвечающих за решение ряда задач, начиная с задач поиска информации и заканчивая задачами управления и распознавания образов. Исследовательские работы в области использования многоагентного подхода проводятся давно, но задача построения универсальной модели многоагентной системы облачных вычислений до сих пор не решена из-за сложности и разнообразия использования аппаратных и программных платформ.

Учеными Национального института стандартов и технологий США разработана эталонная архитектура облачных вычислений состоящая из 5 акторов (NIST)3.

Основным недостатком данной модели является отсутствие агента безопасности отвечающего за безопасное взаимодействие всех акторов системы. В данной модели не осуществляется мониторинг действий облачного пользователя, так же не представлен актор, отвечающий за безопасность облачной системы в целом. Отсутствие данного актора может привести к утечке информации, несанкционированному доступу и искажению информации, использованием злоумышленника информации в

своих целях. Данный недостаток является существенным, а система требует доработки.

В статье О.Ю. Песковой, К.Е. Степовой рассмотрена референтная архитектура облачных вычислений с выделение 5 акторов и представлен ряд требований по обеспечению безопасности облачных сервисов. Авторы акцентируют внимание на ряде организационных и технических требований, но не раскрывают технологию обеспечения безопасности в системе облачных вычислений [1—3].

Авторами ряда статей определены требования к системе защиты информации для виртуальной инфраструктуры, построены модели нарушителя и модели угроз, но в модели виртуальной инфраструктуры отсутствует подсистема, отвечающая за безопасность функционирования системы в целом [4—9].

Опираясь на результаты исследований отечественных и зарубежных ученых в области защиты информации и информационной безопасности систем, таких как П.Д. Зегжда, Д.П. Зегжда, В.С. Заборовский, В.И. Будзко, В.А. Курбатов, Р.Б. Ли, С. Вогля предложена обобщенная структура модели многоагентной системы облачных вычислений, предназначенная для информационной поддержки управленческих решений и защиты системы облачных вычислений, обеспечивающая выполнение ряда функциональных возможностей:

- получения достоверной и объективной информации о поведении пользователя и акторов в системе облачных вычислений для дальнейшего построения стратегии защиты системы;
- сохранение шаблонов пользовательских запросов и действий в системе, необходимых для получения часто используемых данных и выявления нерегламентированных действия и попыток воздействия на систему облачных вычислений;
- использование агента безопасности для мониторинга действий облачного пользователя и акторов в системе;
- расширение функциональных возможностей системы облачных вычислений путем увеличения способов анализа

и интерпретации данных о поведении пользователей полученных из системы, то есть возможность дополнения определенных параметров системы и их взаимосвязей, а также увеличения их количества в процессе использования ресурсов системы облачных вычислений;

- проведение автоматизированных исследований статистических данных о поведении пользователей и акторов в системе, а также минимизации числа ошибок системы, обусловленных человеческим фактором.

В данной статье предлагается модель системы облачных вычислений с выделенным агентом безопасности, он необходим для мониторинга действий облачного пользователя и акторов системы, так же для обеспечения безопасности системы в целом. Агент безопасности строится с использованием автоматной модели.

(from A. Borovki et al. *The Building of a Model of a Protection System in Cloud Technologies Based on the Multi-Agent Approach Using an Automatic Model*)

Practice 15. Write the Introduction to your project proposal (recommended word count is 300). Use the vocabulary of the unit. Also, refer to the Academic Phrasebank by the University of Manchester at http://www.phrasebank.manchester.ac.uk/ or via the QR-code below:

A. Upon writing the introduction, check it with the following checklist:

- My Introduction contains the background which contextualizes the topic of my research.

- My Introduction gives delimitations of the study and points to a gap, or a niche which my research can fill or at least suggest ways of filling it.
- My Introduction contains the aim(s) and/or objective(s) of my research.
- My Introduction explains the professional significance of my study.
- My Introduction is 300 — 400 words long.

B. Swap your Introductions with your partner and assess each other's Introductions against the checklist and the assessment criteria in Appendix C.

C. Go to Appendix G to read an Introduction adapted from the past project proposals. Check it against the checklist. Then assess it according to the assessment criteria in Appendix C.

Unit 3

WHAT MAKES A GOOD LITERATURE REVIEW

Lead in

1. What ways to introduce other people's opinions do you know?
2. What main problem do we face when rendering scientific papers?

Ways to avoid plagiarism:

- Paraphrasing
- Summarizing
- Direct quoting

Practice 1. **Read the following passages and say which one is a) an acceptable summary; b) a plagiarized version; c) a legitimate paraphrase.**

Original excerpt:
Decision trees represent a supervised approach to classification. A decision tree is a simple structure where non-terminal nodes represent tests on one or more attributes and terminal nodes reflect decision outcomes. The ordinary tree consists of one root, branches, nodes (places where branches are divided) and leaves. In the same way the decision tree consists of nodes which stand for circles, the branches stand for segments connecting the nodes. A decision tree is usually drawn from left to right or beginning from the root downwards, so it is easier to draw it. The first node is a root. The end of the chain "root — branch — node-...- node" is called "leaf".