TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI
HANOI UNIVERSITY OF SCIENCE AND TECHNOLOGY

Technical Writing and Presentation
How to write scientific texts in English

SOICT - 2020

Contents

- *Simple rules for writing*
- *Parts of a sentence*
- *Sentence*
- *Combining clauses*
- *Punctuation*

Simple rules for writing

- Have one idea per sentence or paragraph and one topic per section.
- Have a straightforward, logical organization.
- Use short words.
- Use short sentences with simple structure.
- Keep paragraphs short.
- Avoid buzzwords, clichés, and slang.
- Avoid excess, in length or style.
- Omit unnecessary material.
- Be specific, not vague or abstract.

Parts of a sentence

- Skeleton
  - verbs
  - nouns
  - pronouns
  - numerals
  - adjectives
- Additional stuff
  - Adverbs
  - Prepositions
  - Conjunctions.

Verbs

- Two important rules
  - The number of subject determines the number of verb
  - Do not mix inconsistent tenses

- Uses of verbs
  - Number and person
  - Active or passive voice, which person?
  - Tenses
  - Noun syndrom
  - Irregular verbs
  - Notes

Number and person

- When the subject is singular third person (she/he/it), the verb needs suffix -s
- The auxiliary verbs have their own special forms (is, can, has, does).
- When the subject is composed of a singular and a plural noun by ”or” or ”nor”, the verb agrees with the noun that is closer.

Notes

- Be careful with special phrases:

  ”A number of new experiments **were** done” (plural)
  ”Plenty of time **was** spent...” (singular)
  ”A few data points **belong** to cluster *X*” (plural)

- If the number of the subject changes, retain the verb in each clause.

✗The positions in a sequence were changed and the test rerun

✓The positions in the sequence **were** changed, and the test **was** rerun.

Tenses

- Default: the present
- It is possible to combine **perf ect** (has been) (and **future**, will be) if needed, but **not the other tenses**
- Past or present prefect (but not both) when you describe previous research (literature review)
- Past tense to describe the experiments and their results
- Notice: Use ”would” with care! It expresses a conditional action.
  - E.g.”it would appear” → ”it appears”

Active voice vs passive voice

- It is often recommended to use active voice,
- In scientific writing passive voice is sometimes convenient.
- Passive voice allows us to draw the reader's attention to the phenomenon or the event, instead of the actor.

  E.g.

  "The probabilities are updated by Bayes rule",

  "The score is assessed on the basis of the training data."
- Often the purpose determines the voice.

Passive voice

- Usually we want to begin with a familiar word and put the new information in the end.

  E.g. before an equation or a definition, we can say ”The model is defined as follows”
- Do not overuse passive, and do not chain passive expressions.
- Use only one passive per sentence

Passive expressions

- ”It is”
  - E.g. ”It is often recommended [reference] that...”
  - Typical verbs in this expression are: say, suppose, consider, expect.
- “We” can be used as passive.

  E.g. ”In Chapter *X*, we define the basic concepts.”

  should be replaced with

  ”The basic concepts are defined in Chapter *X*.”

Other expressions

- ”You” is sometimes used as passive, especially in manuals. **Don’t use it in scientific text!**
- ”People” when you refer generally to people. **Quite a vague expression**, not recommendable!

- ”There is/are” is used when the real subject (what is somewhere) comes later and we haven’t mentioned it before.

  E.g. ”There was only one outlier in the data set 1” v.s. ”The outlier was in the data set 1.
- ”The verb is nearly always ”be” (sometimes ”exist” or something else)

Person

- **Basic rule**: avoid the first person.
- Referring to yourself: you can talk about ”the **author**”.

  E.g. ”All programs have been implemented by the author”
- Gender-neutral language: when you refer to an unknown user, student, etc. try to use gender-neutral language.
  - The most common way is to say ”she/he” or ”he or she
  - ”One” is neutral, but sounds often awkward.

  E.g. ”The learner can define one’s own learning goals”
- Sometimes you can avoid the problem by using **plural**

Notes

- Do not use **short forms** ”isn’t, can’t, doesn’t”, but ”is not, cannot, does not”.
- Use ”be verb+ing” form when something is currently happening or takes some time.

  E.g. ”Thread 2 can be started in the same time when thread is still running”
- Some verbs require that the following verb is in -ing form

  E.g.”Continue splitting until criterion *X* has been reached.”
- ”Noun syndrom” = use of common verbs {be, do, have, make, ...} + a noun ⇒ Prefer illustrative verbs!

  E.g. ”We can get better understanding...” ⇒ “We can understand better”

Nouns

- Plural forms
- Countable and uncountable nouns
- Differences between British and American English

Plural Forms

- **Irregular plural forms**

Singular form	Plural form	Singular form	Plural form
half	halves	analysis	analyses
life	lives	thesis	theses
axis	axes	parenthesis	parentheses
matrix	matrices	basis	bases
child	children	emphasis	emphases
automaton	automata	series	series
vertex	vertices	medium	media
index	Indices (indexes)	criterion	criteria
appendix	appendices (appendixes)	phenomenon	phenomena

Special cases

- **Data** is originally the plural form of **datum**, but nowadays it is frequently used as a **singular word**.
- The same holds for **hypermedia**.

  E.g.

  ”The data **is** biassed”,

  ”Hypermedia **offers** a new way to implement learning environments”

Notes

- If the suffix is *{*-s,-ss,-sh,-ch,-x,-z*}* in singular → -es in plural,
  E.g. research – researches, approach – approaches, quiz – quizzes
- The same happens with most words which have suffix -o, unless the word is abbreviated or of foreign origin.
  E.g.cargo – cargoes,
- but photo – photos, dynamo – dynamos
- After **consonant** -y changes to -ies in plural.
  E.g. floppy – floppies.
- **Singular words which look like plural forms**
  The names of disciplines: mathematics, statistics, physics.
  ”Statistics is the precessor of data mining.”

Uncountable nouns

- Words expressing material: water, air, wood, ...
- Abstract words: life, time, work, strength, ...
- Exceptional: advice, information, news, equipment, money (Plural meaning)

Compound words

- The practices vary, possibly words are written together, with a hyphen -, or separately.
- If the words have become one concept, they are usually written together, e.g. ”software”, ”keyboard”, ”database”
- If the independent meaning of words is emphasized, they are hyphened,
  e.g. ”non-connected”
- Hyphen is often used when the concept consists of more than two words:
  ”depth-first search”, ”between-cluster variation”, ”feed-forward neural network”, ”first-order logic”
- Multiple word adjectives are usually hyphened, e.g. ”data-driven”, ”model-based”, ”class-conditional”
- If the first part is a symbol or an abbreviation, the word is hyphened,
  e.g. ”*NP*-complete”, ”*k*-nearest neighbour method”, ”3-dimensional”

Articles

- Position
- Definite and indefinite concepts
- Basic rules of using articles

Position

- Basic rule: before the noun phrase
- Exceptions:
  - {what, such, quite, rather, half} + a/an + noun phrase

    E.g. ”Too great a distance”, ”so long a time”
  - {too, as, so, how, however} + adj. + a/an + noun

    E.g. ”as big a difference”
  - {all, both, double, twice, half} + the + noun

    E.g. ”All the methods”

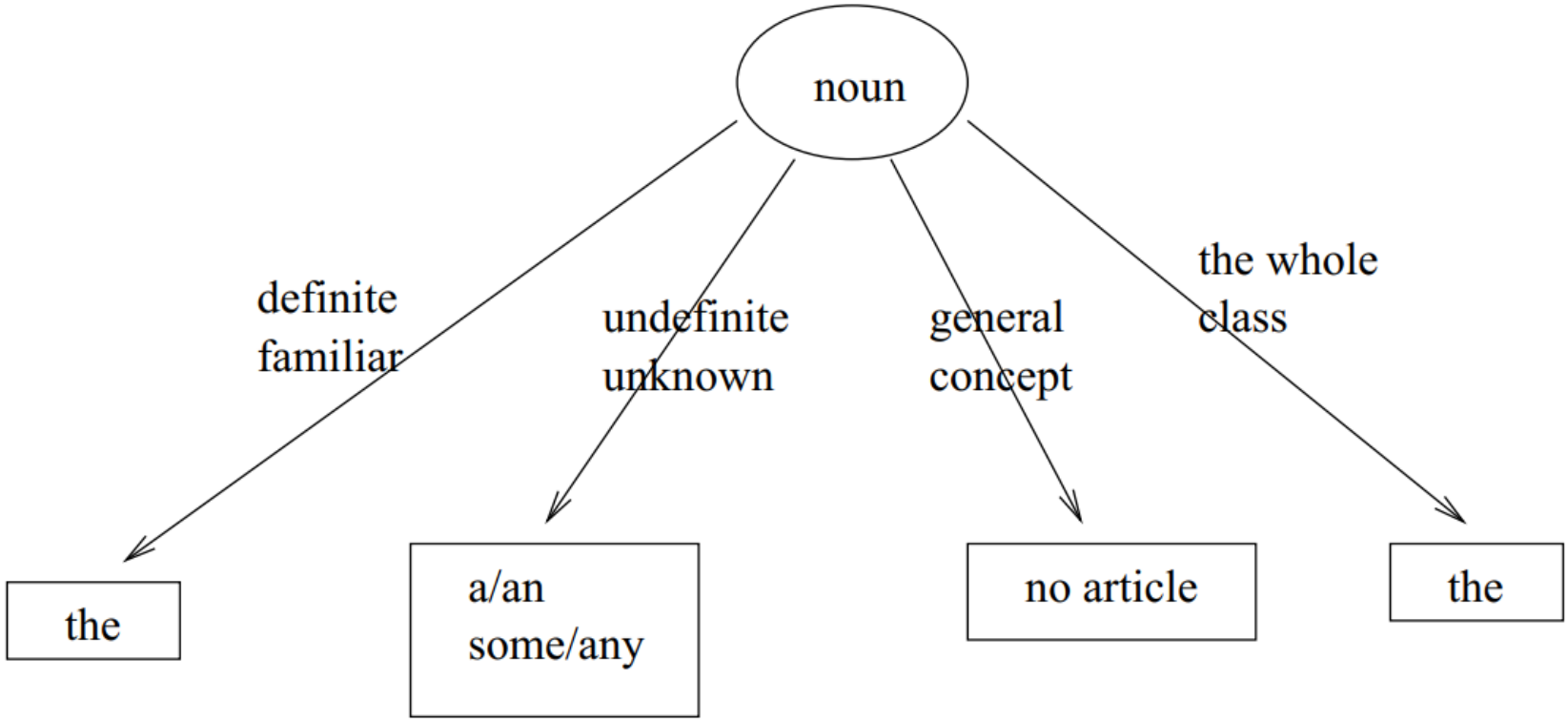
Definite and indefinite concepts

- A concept is **indefinite**, when you mention it first time, and it is not clear from the context

  E.g.”There was a time delay between processes *A* and *B*.”

- It is **definite**, when
  - you mention it again
  - the context defines what you mean
  - the concept is familiar to everybody
  - Usually this kind of expressions are defining: ”The delay between two processes *P*1 and *P*2 is *tend*(*P*1) - *tstart*(*P*2).”

Basic rules of using articles



Refer to an indefinite concept

- a singular countable noun → **a/an**
- a plural countable noun + positive clause → **some**
- a plural countable noun + negative or interrogative clause → **any**
- a uncountable noun + pos. clause → **some**
- a uncountable noun + neg. or interr. clause → **any**

Notes

- **When you refer to something generally**
  a plural countable noun or a uncountable noun → no article
  ”Students need time to process new inforsmation”
- **When you refer to the whole class**
  a singular countable noun → a/an
  ”The computer cannot solve all problems”
- **Exceptional expressions**
  Sometimes you can use a/an article with an abstract word:
  - when the word is proceeded by a describing relative clause ”There is a danger that the model overfits”
  - expressions ”a /short/long time”, ”a while”

"the" article with ordinal numbers and some adjectives

- Definite article "the" is used
  - when the noun is preceded by an ordinal number

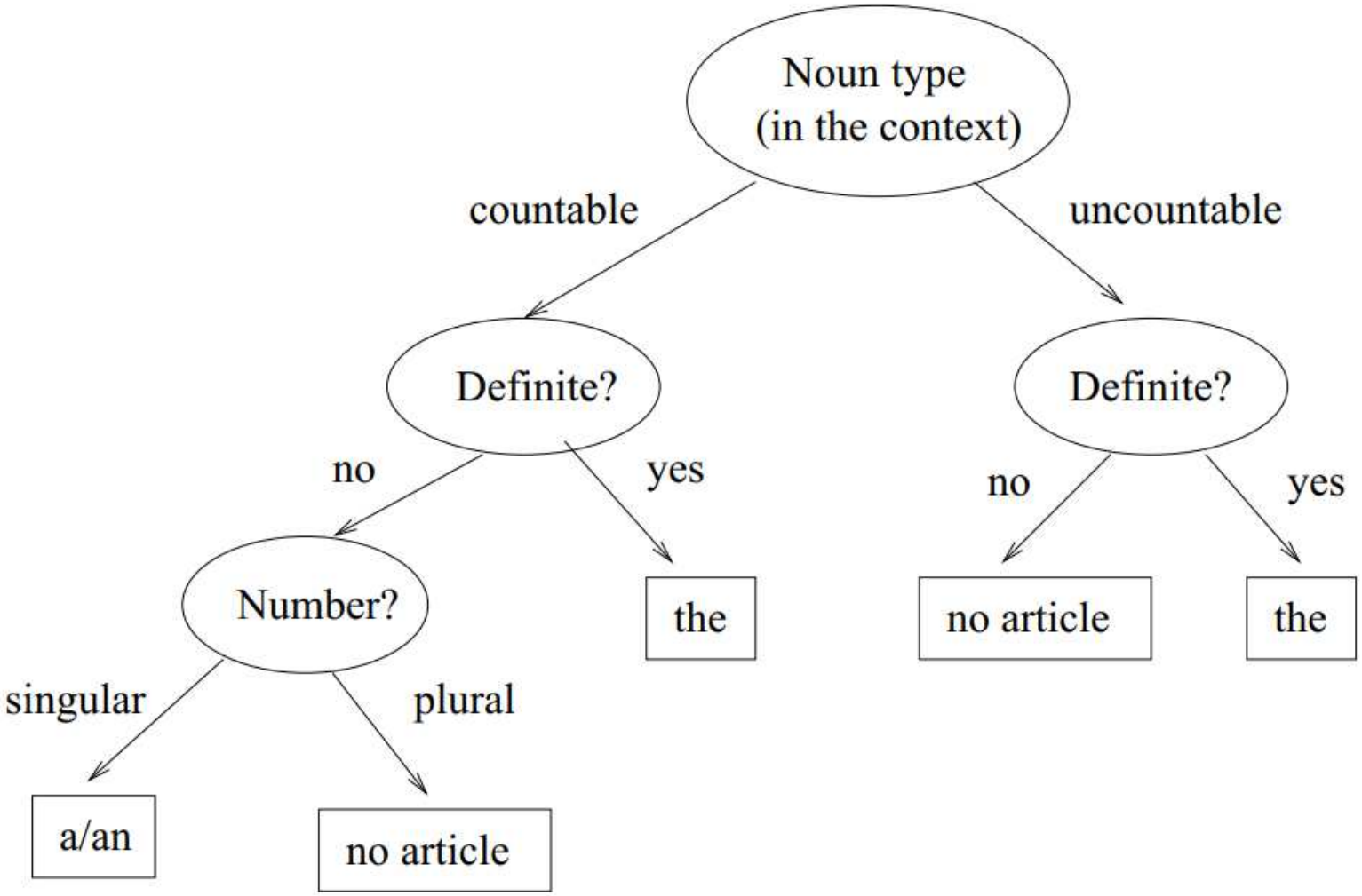
    "The **first** attribute describes..."
  - when the noun is preceded by an adjective expressing order

    "the **next** attribute", "in the **following** chapter"
  - with adjectives **same, only, right, wrong**

    "The results were **the same**", "**The only** model which has this property is *X*"
- Notice: "the" is not used with ordinal numbers or adjective "last", when you refer to the performance in a competition

  "Program *X* came **first** and program *Y* was **last** when the programs were compared by the *Z* test.

Decision tree for articles



Notes

- A noun can be used as a countable or an uncountable concept
  ”It takes time...”, ”all the time”, ”one at a time”, ”many times”
- **any** or **some** before the noun, → indefinite” ,
  E.g.”The disk contains (some) space for back-up files”
- Don’t use pronouns!”This *x*” → ”the *x*” (where *x* is a noun)
- $\exists x \in C$ such that $P(x)$, use article **a/an**. ”
- $\forall x \in C, P(x)$, use article **the.** (only for countable concepts!
- A name without any modifying word → no article,
  E.g.”*X* is independent from *Y* “
- A modifying word like ”set”, vector”, ”model” etc. stands before the name → Two habits:
  1. No article when you mention the entity for the first time. After that use definite article ”the”, or
  2. Never any articles.

Pronouns

- Two important rules when you use pronouns:

  1. When a pronoun refers to a noun in the preceding sentence, make sure that the **referred is obvious**!
  2. Each pronoun should agree with the referent in number and gender.

Notes

- **Pronouns which require singular verb form**
  *{*everybody, anybody, nobody, everyone, anyone, no one*}*
- Every + singular noun
- All + singular or plural noun
- several < many
  several ≈ some

Adjectives

- Two notes:
  1. **Avoid vague adjectives!**
  2. How to derive and use comparative and superlative forms?

use the comparative

- Basic structure:
  - *X* is **as** efficient **as** *Y* (*X* and *Y* are equally efficient)
  - *X* is more efficient **than** *Y*
- Exceptional expressions:
  - *X* is **different from** *Y*
  - *X* is **similar to** *Y*
  - *X* is **the same as** *Y*
  - *X* is **inferior/superior to** *Y*
  - *X* is **equal to** *Y*

Adverbs

- Use as few adverbs/prepositional phrases as possible!

→ expressive verbs and nouns

- Use introductory adverbs like ”fortunately, similarly, conversely, certainly” carefully
- Notice that ”importantly” and ”interestingly” are not proper adverbs.
  E.g.
  ”More importantly, the accuracy can actually increse when the complexity is reduced”
  → ”More important, the accuracy can actually increase when the complexity is reduced.”
  ”Interestingly, we found that...”
  → ”An interesting finding was that...”

Special cases of adverbs

still and **yet**

- Still: before the main verb, but after be-verb.

  ”These enlargements are still unimplemented”

- Yet: at the end.

  ”These enlargements have not been implemented yet.

so and **such**

- So: before adjectives or adverbs which are **not** succeeded by nouns

  E.g. ”The time complexity is not so hard”

- Such: when an adjective is succeeded by a noun.

  E.g. ”Such time complexity is infeasible”

  - Notice the article ”a/an”, if the noun is countable:

    ”such a system”

    ”such an algorithm

Prepositions

- Be careful with prepositions. A wrong preposition can give a totally different meaning!
  • Hint: When you use a preposition, visualize the direction it is signaling and ask yourself if it is appropriate.
  • If you are unsure about the use of a preposition, ask yourself what a cat would do! (Fedor's sciwri book)
  Cats sit **on** mats, go **into** rooms, are part **of** the family, roam **among** the flowers

Expressing location

- Usually **in**, e.g. ”in set *X*”
- If an exact location, then **at**, e.g. ”at point (*x, y*)”
- If the location can be imagined as a line or a surface, then **on** ”on the *x*-axis”, ”on a time line”
- Special cases:
  - ”**on** page 3”, ”**on** line 5”, ”**on** the Internet”
  - ”A file is loaded **from** the hard disk **into** main memory.”
  - ”results **from** the survey suggest...”
- over – under/beneath
- above – below

  ”*X*’s points were **below** the average points”
  ”The task is to optimize *f* **under** the given constraints”

Expressing time

- **Exact time**: **at**, e.g. ”at the moment”, ”at four o’clock”, ”at the same time”
- **Longer period of time**: **in**, e.g. ”in the 1970’s”, ”in the future”, ”in five minutes”,
- Notice: ”In the beginning/end” vs. ”At the beginning/end” of something

Expressing the target or the receiver

- **to or for?**
- Basic rules:
- When direct receiver, then **to**
  ”The values are assigned to variables”
- When the final receiver (for whom something is meant) then **for**
  ”I gave the book for Belinda to Tersia”
  ”The messages for nodes *F* and *G* are transfered to node *D* for rerouting”
  ”A variety of methods have been developed for the clustering problem.”
  • When something is good or bad for something, then **for**
  ”Problem-based learning is good for students”

Expressing the target or the receiver

- Some verbs require either for or to:
  1. If the verb is *{*bring, give, take, show, offer*}* → **to**
  2. If the verb is *{*be, get, keep, make*}* → **for**
- Sometimes the preposition can be missing, depending on the word order:
  i) verb + receiver + object
  ii) verb + object + to/for + receiver
  iii) verb + to/for + receiver (no object)

Special phrases

- constraint **on** something
- independent **from** something but dependent **on** something
- different **from** something but similar **to** something
- difference **between** something and something
- prefer something **to** something
- impact of something **on** something
  influence **on** something
  effect **on** something (but to affect something)
  focus **on** something
  perform **in** some manner
  generalize **to** something
  a discussion **about/on** something
- research **on** something but a study **of** something
  reason **for** something
- opportunity **of/for** something

in spite of something (but despite something)
regardless of something
take into account
in relation to something
in contrast with something
a proportion **of** something. (“a large proportion of data”)
in proportion **to** something, proportional **to** something (“The time complexity of *f* proportional to *n* is...”)
the ratio **of** *a* **to** *b* = *a*/*b*
x% **of** *y*
under some conditions
by default
contrary to something
in contrast
by contrast (~ “however”)
on the contrary
at an extreme

Parallel structures

- **Parallel structure** = words, phrases, clauses or sentences combined by commas and/or conjunctions.
- Here we call the combined items as **parallel items**.
- Parallel items are combined by parallel conjunctions (and, or, but, ...).
- lists are also parallel structures!
- Often the parallel structure lists alternatives or makes some kind of comparison: the items belong to the same or similar classes or to two opposite classes.

  E.g. ”Method *X* has several advantages: it is easy to implement, it works in polynomial time, and it can use both numeric and categorial data.”
  contains two parallel structures: three advantages (”it is, it works, it can”) in a list and ”both numeric and categorial data”

Basic rules for parallel structure

- The parallel structure should be consistent in two ways
- **Semantically**: the concepts referred by parallel items should be comparable
- **Syntactically**: the items should have similar grammatic structure. All of them should be either nouns, noun phrases, verb phrases, or clauses.
- Parallel items should be in the same form, e.g. you cannot combine ”to” + verb and a verb without ”to”.

  E.g. ”The problem is both hard to define and solve”
  → ”The problem is both hard to define and **to** solve

conjunction pairs

- Parallel items combined by conjunction pairs
  - between...and,
  - both...and,
  - either...or,
  - neither...nor
  - not only...but…(also)

”*X* solves the problems of traditional clustering algorithms. **Neither** outliers **nor** missing values affect the clustering quality.”
”The task is **not only** easy to solve **but** it can **also** be solved efficiently

The comparative – the comparative

- the + comparative + *x* + comma + the + comparative + *y*, where *x* and *y* complete the clauses
  ”**The more complex** the model is, **the better** it describes the training data.”

Parallel sentences

- The sentences in the list begin by ordinal numbers ”First, Second, Third”
- ”*X* model has three important properties: **First**, the model structure is easy to understand. This is a critical feature in adaptive learning environments, as we have noted before. **Second**, the model can be learnt efficiently from data. There are feasible algorithms for both numeric and categorial data.**Third**, the model tolerates noise and missing values.”

Sentences

- Terminology
- Sentence styles
- Types of dependent clauses
- Sentence length
- Word order
- Combining clauses

Terminology

- A **sentence** consist of one or more clauses
- A **clause** contains always a subject and a predicate, and usually an object
  – An **independent clause** (main clause) can make a sentence alone.
  – A **dependent clause** (subordinate clause) needs an independent clause for support

Sentence types

- The sentence type depends on the type of its main clause. The main types are following:
  1. Statement (ends by a full stop: ”*x* is *y*.”)
  2. Question (ends by a question-mark: Is *x y*?”)
  3. Order (ends by an exclamation mark: ”Be *x y*!)
- In scientific writing the default type is the statement.
- Direct questions and orders are seldom used.
- Questions suit best to the introduction where you state your main research questions clearly and concretely

Types of dependent clauses

1. Clauses beginning by sub-ordinating conjunctions (when, if, because,while, …)
2. Relative clauses (begin by relative pronouns which, who, that)
3. Indirect questions (begin by question words or if/whether)

Examples:

”The dependency is trivial, **because** $Y = f(X)$.”
”*X* and *Y* are linearly independent, **if** the correlation coefficient, *corr* (*X, Y*), is zero”
”Let *ci* be the cluster **which** is closest to *x*.
”We select the first model **that** fits the data.”
”First we should study **what** is the relationship between *X* and *Y* ”
.
”The main problem is **whether** *X* can be applied in *Z*.”
”We analyze the conditions **under which** *X* can be applied.”

Sentence length

- always less than 30 words, preferably less than 20 words!
- 1-3 clauses
- expresses one idea
- If you tend to write too long sentences, try the following:
  - Identify the main subject-predicate-object section
  - Prune or compress everything else, which is not needed
  - Check the verb structures and ask yourself if they could be shorter

Word order

- The order of words has a strong impact on the meaning!

  E.g.

  ”There is, however, currently no information about the limitations of quantum computers.” →

  ”However, there is no current information about the limitations of analog computers.” →

  ”However, the limitations of current quantum computers are not known.”

Word Order

- Look at 3 examples:

  ”There is, however, currently no information about the limitations of quantum computers.”
  ”However, there is no current information about the limitations of analog computers.”
  ”However, the limitations of current quantum computers are not known.”

Word order in different languages

SOV	Urdu, Ancient Greek, Bengali, Hindi, Japanese, Korean, Latin, Persian, Sanskrit
SVO	Chinese, English, French, Hausa, Italian, Malay, Russian, Spanish, Thai
VSO	Biblical Hebrew, Arabic, Irish, Filipino, Tuareg-Berber, Welsh
VOS	Malagasy, Baure, Car
OVS	Apalaí, Hixkaryana, Klingon
OSV	Warao

Basic Word Order

- **subject-predicate-object** (SVO)
- You can add attributes, phrases and clauses, but don't deviate too far from the basic format
- Goal: put the most important information to the beginning of a sentence or begin by a familiar thing and put the new information to the end
- E.g. "*X* is a new algorithm for the TS problem" "The probabilities are updated by the Bayes rule:" + the equation.
- Often the sentence is most informative, if you express the most important topic by the subject.
- This format helps to write clear and compact sentences

More details about word order

- The adverbs and prepositional phrases occur in order: **way, place, time**
  - ”The nearest neighbours can be identified **efficiently** (way) **in a dendrogram** (place)”.
  - ”The values can be updated **easily** (way) **in linear time** (time)”.
    Verb modifiers: in the middle of clause
- Adverbs which express frequency: always, ever, never, often, seldom, sometimes, usually
- **Adverbs which can begin the clause**
  If the adverb expresses time, it can be also in the beginning:
  ”Next, the data is loaded to the main memory.”
  This gives more emphasis to the word. It is also used, when there are other adverbs/prepositional phrases in the end of the clause.
  Introductory adverbs like ”obviously”, ”fortunately”, etc. are always set to the beginning (if they are needed)

Combining clauses

- Say the main message in the independent clauses! Use dependent clauses only to add details
- Contents
  - Combining two independent clauses
  - Combining clauses by sub-ordinating conjunctions
  - Relative clauses
  - Indirect questions

Combining clauses by sub-ordinating conjunctions

- The basic form: an independent clause + a sub-ordinating conjunction + a dependent clause.
- The most common sub-ordinating conjunctions express

1. a chronological order: **when, as, as soon as, while, after, before,until, since**

   ”The search can be halted as soon as *minfr* proportion of data is checked”

2. a conditional relationship: **if, unless**. **If**-clauses can also begin the sentence

   ”If the order is fixed, the episode is called serial.”

3. a reason: **because** (Recommendation: reserve word ”since” to express chronological order)

   ”The method is time-efficient, because all the parameters can be updated in one loop

4. a purpose: **so that** (You can also use **in order to** + infinitive verb)
5. an admission: **although, even if**

Notes

- When you combine
  - an independent clause + a dependent clause → sometimes but not always a comma (e.g. before **but**, but not before **that**).
  - a dependent clause + an independent clause → always a comma

Relative clauses

- Relative clauses can be divided into two categories: **Restrictive** or **Non-restrictive**
- **who** refers to a person.**whose** is the genitive form, it can refer to things and objects, too!
- **whom** is used as an object and with prepositions,
- **which** refers to things and objects, can be used as a subject or an object or with prepositions.
- In most cases, the genitive form can be either **of which** or **whose**.
- **that**
  - can be used only in restrictive relative clauses! → never use comma before it!
  - can refer to people or things,
  - can be used both as a subject and an object.
- If you need prepositions, they have to be in the end of the clause!
- **what** contains also the correlate

Combining two independent clauses

- A **compound sentence**= two or more independent clauses which are combined by co-ordinating conjunctions or (rarely) by semicolons.
- Combine only two main clauses (unless the clauses have the same subject which is mentioned only once).
- The ideas expressed in the clauses must be closely connected (otherwise separate sentences).
- The most common co-ordinating conjunctions are **and** and **but**.
  - **and** links one idea to another.
  - E.g. ”The data is sparse and the model overfits easily.”
    but establishes an interesting relationship between the ideas → a higher level of argument.
  - E.g. ”The data was sparse, but the model did not overfit.”
    (=”Even if the data was sparse, the modeldid not overfit.”)

Indirect questions

- The dependent clause begins by a question word **what, why, when, where, how** or **if/whether** when the corresponding direct question begins by a verb.
  ”First we should study **what** is the relationship between *X* and *Y* .”
  ”The main problem is **whether** *X* can be applied in *Z*.”
  • The word order is direct!
  • No auxiliary word **do**
  • No comma!
  • No question mark

Punctuation

- **Goal**: to make the text clearer.
- The English punctuation rules do not always coincide with the rules of your mother tongue.
- Usually you manage with just two marks: **full-stop** and **comma**!
- The basic rules for other marks are:
  - Use colon ':' only when needed.
  - Avoid semicolon ';' and dash '–'.
  - Avoid unnecessary parantheses '('...')'.

Comma is used

1. To separate introductory phrases and conjunctions (however, thus, similarly, etc.)

”Despite the high time complexity, *X* is often used...”

2. When the sentence begins with a dependent clause.

”Since *x* is a statistic, it is also a random variable.”

3. When a non-restrcitive relative clause is embedded into an independent clause or ends a sentence.

”*X*, which is responsible for data preprocessing, initializes *Y* .”

4. When two phrases with the same meaning are used side by side.

”One of the most useful statistics is ***x*, the sample mean**.”

5. When the sentence begins by an infinitive structure

”To find the lower bound for the confidence interval, we isolate...”

6. To separate items in a list of three or more items.

7. To avoid ambiguity.

”Instead of hundreds, thousands rows of data is required”

No comma is used

- When an independent clause is followed by a restrictive relative clause or is embedded with a restrictive rel. clause (especially before **that**). Exception: ”It must be remembered, however, that...”
- Between two independent clauses (in British English).
- Before an indirect question.
- When you begin by a prepositional phrase expressing the place.
- ”In this section we discuss...” ”In Chapter 3 we defined...”