

12 Sentences

12.1 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.

12.2 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, e.g.

"The main research questions are the following:

1. What is the relationship between X and Y ?
2. When X can be applied?
3. Can we apply X in Z ?
4. How X can be extended?"

Orders can be useful in pseudo code, when you describe some method. E.g. "Search such c_i that $d(x, c_i)$ is minimal".

Dependent clauses can be divided into the following types:

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 c_i 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."

12.3 Sentence length?

Recommendations:

- 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:

1. Identify the main subject-predicate-object section
2. Prune or compress everything else, which is not needed
3. Check the verb structures and ask yourself if they could be shorter

E.g. verb structure "has been shown" can often be replaced by "is".

Notice! Don't go into the other extreme when you shorten sentences! If the clarity suffers, then a longer sentence is better.

Analogue: A good model of data does not overfit nor underfit, i.e. it is simple enough but still expresses all essential features. Now the sentence is a model of the idea you want to express.

12.4 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."

12.4.1 The basic word order: subject–predicate–object

Recommendation: use the basic format **subject-predicate-object** in your sentences. You can add attributes, phrases and clauses, but don't deviate too far from the basic format.

Why?

- Goal: put the most important information to the beginning of a sentence! "X is a new algorithm for the TS problem"
- Or begin by a familiar thing and put the new information to the end "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

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)".

12.4.2 Verb modifiers: in the middle of clause

Words which modify the predicate (the main verb) are located in the middle of the clause:

- Adverbs which express frequency: always, ever, never, often, seldom, sometimes, usually.

- Adverbs which express degree: almost, quite, certainly, completely, hardly, just, only, quite, really.
- Other words which modify the verb: already, also, still.

Hint: always consider if the word modifies the verb (the action) or the object (the target).

Verb modifiers are located

- before the predicate, if the verb consists of one word and is not the "be"-verb.
- after the first auxiliary verb, if the verb consists of several words.
- always after the "be"-verb.

E.g.

"X often implies Y."

"The method gets sometimes stuck at a local optimum"

"The data was probably biased."

Task: Draw a decision tree for deciding the position of adverbs!

Problem: some words like "only" can modify also other words!

→ Put the word "only" next to the word or phrase it modifies!

E.g. (notice the different meaning):

"X was the only method which could parse the *LL*(1) grammar"

"X was the method which could only parse the *LL*(1) grammar"

"X was the method which could parse only the *LL*(1) grammar"

12.4.3 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).

12.5 Combining clauses

Say the main message in the independent clauses! Use dependent clauses only to add details.

12.5.1 Combining two independent clauses

A **compound sentence**= two or more independent clauses which are combined by co-ordinating conjunctions or (rarely) by semicolons.

- In principle, you can combine several independent clauses, but in practice, 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** just links one idea to another (doesn't describe the relationship – typical for the children's language and dreams where things just happen). 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 model did not overfit.")
- Commas? If the clauses have the same subject, then no commas. Otherwise usually a comma, unless the clauses are very short.

12.6 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**

2. a conditional relationship: **if, unless**. **If**-clauses can also begin the sentence: "If the order is fixed, the episode is called serial." Notice: unless = if... not
3. a reason: **because** (Recommendation: reserve word "since" to express chronological order)
4. a purpose: **so that** (You can also use **in order to** + infinitive verb.)
5. an admission: **although, even if**

Examples:

"The search can be halted as soon as min_{fr} proportion of data is checked"

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

When you combine

- an independent clause + a dependent clause \rightarrow sometimes but not always a comma (e.g. before **but**, but not before **that**).
- a dependent clause + an independent clause \rightarrow always a comma.

12.7 Relative clauses

Correlate = referred word or clause, e.g. "An outlier is a **data point** which lies outside the clusters", "**Students** who solved the task...", "**The time complexity is quadratic**, which is sometimes undesirable..."

12.7.1 Restrictive and non-restrictive relative clauses

Relative clauses can be divided into two categories:

1. **Restrictive** or essential relative clause
 - defines the correlate
 - is necessary for understanding the sentence correctly
 - no commas
 - "X is an algorithm which solves the Travelling Salesman problem in $O(n^k)$ time."

2. **Non-restrictive** or non-essential relative clause

- gives only additional information about the correlate
- is separated from the main clause by commas
- "X, which solves the *TS* problem, works in $O(n^k)$ time."
- can refer to the previous clause

Notice that in spoken language we can drop the relative pronoun, if it is not the subject of the clause. In scientific writing, it is better to write all pronouns, because it should always be clear what you are referring.

12.7.2 Which relative pronoun to select?

1. **Who, whose, whom**

- **who** refers to a person.
- **whose** is the genitive form, e.g. "The student, **whose** solution was correct, got extra points."
- Notice that **whose** can refer to things and objects, too!
"X is an example of problems **which** belong to class *NP* and **whose** known solutions are exponential."
- **whom** is used as an object and with prepositions, e.g. "The student, **about whom** I told you yesterday, wants to speak to you."

2. **which**

- refers to things and objects.
- can be used as a subject or an object or with prepositions.
"X is a trick **which** helps to estimate the parameters more accurately."
"Let X be the variable **which** Y depends **on**."
- when you refer to an entire clause, e.g. "The time complexity is quadratic, **which** is sometimes undesirable...", "All students cannot study themselves, **which** means that tutors are needed."
- In most cases, the genitive form can be either **of which** or **whose**. E.g. "A computer, the cache **of which** is disabled, is less efficient..." = "A computer, **whose** cache is disabled, ..."

- However, when you refer to abstract nouns, use **of which** structure:
"N.N. has introduced a new method the complexity **of which** is exponential." (**Problem:** is this rule still valid? Could we nowadays use also "whose"?)

3. that

- **can be used only in restrictive relative clauses!** → never use comma before it!
- can refer to people or things,
"The student that has solved the task", "The task that was solved"
- can be used both as a subject and an object.
- If you need prepositions, they have to be in the end of the clause!
"The problem **that** we talked **about**..."
- **that** is often used, when
 - the correlate is {**all, little, much**}, e.g. "all that we know"
 - with **superlative**, "the best solution that we can invent"
 - with ordinal numbers (**first, second,...**), **last**, and **only**,
"the only algorithm that is comparable with X is Y "

4. what

- **what** contains also the correlate
- Only in special expressions!
- E.g.
"WYSIWYG means 'What you see is what you get'."
"This is what we know so far."

12.7.3 Extra: When to use who and when whom?

Problem: Is the relative pronoun a subject or an object (who or whom)?

Hint: turn the subordinate clause around and substitute the relative pronoun by a personal pronoun. If you can use "she" or "he", it is subject

(who), if "her" or "him", then it is an object (whom).

E.g. "N.N. is the student who/whom I mentioned earlier"

→ "N.N. is the student. I mentioned her/him earlier"

→ "whom" is the correct choice.

12.8 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