

A large red square with a white border, centered on a white background. Inside the square, the text "Guidelines for Annotators: Discourse Questions" is written in white, bold, sans-serif font, centered horizontally and vertically.

Guidelines for Annotators: Discourse Questions

Task Overview

- We wish to extract how statements, e.g. events, relate to each other.
- We want to extract these discourse relations by relating two events through a question and answer pair.
- Events can be related in various ways, for example temporally or causally.

The 22-year-old victim, who has not been identified, **was sitting inside the vehicle** and **also suffered numerous gunshot wounds**.

- In this sentence we have, for example, the event of **someone sitting inside a vehicle** and the event of **someone suffering gunshot wounds**.
- Here the most natural relation between the two is a temporal one, e.g. they both happen at the same time. This could result in the following QA pair:
 - Q: While what did the 22-year-old victim suffer numerous gunshot wounds?
 - A: Sitting inside the vehicle

More examples:

Despite labor-shortage **warnings**, 80% **aim** for first-year wage increases of under 4%; and 77% say they'd **try** to replace workers, if **struck**, or would **consider** it.

- Let's say you are looking at the following two events: **trying to replace workers** and **being struck**.
 - The relation between the two is a *conditional relation*, thus you can form this question and answer:
 - Q: In what case would they try to replace workers?
 - A: If they are struck
- Let's say you are looking at the following two events: **trying to replace workers** and **considering it**.
 - The relation between the two is that *either of the two might be a possibility*, thus you can create this question and answer pair:
 - Q: What is an alternative to them trying to replace workers?
 - A: Or they would consider it

What you have to do:

- Read the sentence and check the event(s)/statements you are targeting, they will be marked.
- Think of what kind of relations the events/statements in the sentence express.
- Ask a question using our question templates.
- Answer your question by highlighting a span of the sentence, so that the answer contains the other event.

Annotation Interface

Not having **shown** interest in religion before, he **became** interested in religious thought.

Ask a question where the **highlighted word** occurs either in the **question** or in the **answer**. We marked other potential targets in bold as hints for you.

Question

(Question Start)  (Auxiliary)  Mark a part of the text  ?

The Question

Question: Question Start

We created a list of Question Starts for you. You need to choose one of those to begin your question. They can serve as hints for the type of questions we are looking for.

Question starts:

After what
Before what
Until when
While what

Temporal

What is the reason
What is the result
In what case
Unless what

Causal

Despite what
What is contrasted with
What is similar to

Comparison

Instead of what
What is an alternative to
What is an exception of
What is an example of
In what manner

Other

Question: Body

Next, given the question start, you need to form the complete question, which involves 3 steps:

1. Choose an auxiliary verb if needed.
2. Mark and thus copy text from the sentence.
3. Modify the copied text from the sentence, if required, to make the question grammatical.

Question: The Auxiliary

Depending on the question start and on the content of the question, you might need to select an auxiliary (did, is, might, will etc.).

For example:

- The passengers died after impact with the sea.
 - After what **did** the passengers die? - after impact with the sea
- Erin is helping to design a lab exercise for college students around the US to do the same.
 - For what purpose **is** Erin helping to design a lab exercise? - for college students around the US to do the same

Question: The Auxiliary

While some types of questions require you to select an auxiliary to form a grammatical question, others don't. In those cases you can select the None option.

- For example:
 - Some materials were salvaged from the old dam to build the new, and some have been put on display.
 - What is the reason some materials were salvaged from the old dam? - to build the new dam
 - What is an alternative to some materials being salvaged from the old dam? - some materials have been put on display

Question Body

- Now you need to actually select the **main content** of the question.
- Try to **copy** as much as you can from the sentence, instead of writing the words yourself.
- You are allowed to copy **multiple spans** from various parts of the sentence.
- You need to make sure that you are asking a question about an “eventive” target. Usually a verb is considered “an event”, but there are also cases when nouns or other words, can be events, for example “hurricane”.
- There are the following main issues you need to consider, which we will explain in the following slides:
 - a. Grammaticality
 - b. Span

Question Body: Grammaticality

- We want our questions to sound as natural and grammatical as possible.
- Since you are copying words from a sentence into a question, you will usually have to make modifications.
- Most of these modifications are simple, e.g. switching from “became” to “become” or from “I” to “me” etc.

Sentence: It is unclear if the suspects left with any property, but early reports indicate one suspect was injured.

Question: Despite what is it unclear if the suspects left with any property?

Here, because you added the auxiliary “is”, you need to delete the “is” after the “it”.

Question Body: Grammaticality

- As a way to help you formulate the questions, try, if possible, to stick to the following format:
 - Question Prefix + Auxiliary + Subject + Verb + Additional Arguments

Sentence: You could either go to sleep or **play** a bit longer on your phone.

Question: [What is an alternative to (question prefix)] [(no auxiliary)] [**you** (subject)] [**playing** (verb)] [**a bit longer on your phone** (additional argument(s))]?]

The example shows how you might need to copy from different parts of the sentence for a complete question. While it might have been easier to simply ask “What is an alternative to playing a bit longer on your phone?”, we want you to stay consistent and also include the subject “you”.

Question Body: Grammaticality: Placeholders

Sentence: Many people attended the concert, which was surprising.

Question: What is the reason something was surprising?

In some cases you might have to introduce words into the question that don't occur in the sentence. We call them **placeholders**.

The allowed placeholders are: **someone**, **something**, **to do** and **for**. In the example above, the **something** had to be added in order to make the question grammatical.

Question Body: Span

The span concerns how much you should include from the sentence to form the question.

In essence we want you to keep the span as **concise** as possible, while at the same time **including all the semantic information pertaining to the specific question and answer** you are asking.

Additionally, when in doubt about what span to include, a good rule to keep in mind is that the span you are including in your question **should only contain a single event**. (this is just a rule of thumb when in doubt, sometimes there are cases when you need to include multiple events in the span to make the question natural!)

This is best illustrated with examples (see following slides)

Question Body: Span - Example

Sentence: Passengers were seen using emergency slides to escape the wreckage.

Question: What is the reason passengers were using emergency slides?

Answer: to escape the wreckage

Note that “seen” was not copied into the question! There is a subtle but distinct **difference in meaning** when asking “What is the reason passengers were seen using emergency slides?” vs. “What is the reason passengers were using emergency slides?”.

And the answer “to escape the wreckage” only **makes sense** with the correct question.

Question Body: Span - Example

Sentence: By August that year, the committee had received a recommendation from Mr. Gran Johnson of the Australian Forests Products Association for the “re-introduction of harvesting activities in forest areas previously set aside for conservation.”

Question: While what has the committee received a recommendation?

Answer: By August that year

Here it would have been **too much to include the whole span** after “the committee had received a recommendation”, e.g. “While what has the committee received a recommendation from Mr. Gran Johnson of the Australian Forests Products Association for the ... ?”.

The essential events that you are relating here are “August” and “received”, all the additional context is not needed to form a meaningful QA pair.

Question Body: Span - Example

Sentence: Half of the fourteen passengers and two crew died “instantaneously” and the rest died very shortly after impact with the sea from blunt force trauma, the inquiry found.

Question: After what did the rest die?

Answer: Half of the fourteen passengers and two crew died “instantaneously”.

Here it would have been **too much to include the whole span** after the span included in the question. The question you are asking is comparing the event of the “the rest dying” vs. the event of “dying instantaneously”, telling us that one of the events happened before the other event.

Note that we didn’t include “from blunt force trauma” and “after impact with the sea”, **since they each represent an event in itself**, e.g. you could for example ask the question “What was the reason the rest died very shortly after impact with the sea?” - “from blunt force trauma”.

Question Body: Span - Example

Sentence: When Wall Street turned, however, the big boys entered the market, looking for bargains.

Question: After what did the big boys enter the market, looking for bargains?

Answer: When Wall Street turned

In this example we chose a larger span for the question, also including “looking for bargains”, because they both happen together after “Wall Street turned” and therefore this span contains information relevant to the QA pair and the Question Start.

Question: Multiple Questions/No Question

Sometimes you can ask multiple questions about the same two events (see example in positive examples).

And sometimes none of the question options are valid, in which case you should select “No Question applicable”.

The Answer

Answer

Selecting the answer to your question should be straightforward: select the span that most precisely answers your question.

There are some additional things to keep in mind:

1. Try to stay natural and grammatical.
2. Follow the span guidelines.
3. Include the connective/trigger.

Answer: Special Cases: Connective/Trigger

Sentence: They wanted a court to block today's scheduled liftoff long enough for them to ask the World Court to order a permanent cancellation of the \$1.5 billion flight.

Question: Until when do they want a court to block today's scheduled liftoff?

Answer: long enough for them to ask the World Court to order a permanent cancellation of the \$1.5 billion flight

Many times the trigger won't be one of the typical connectives that you might expect and instead the relation might be expressed with other words. Please also include these cases, as seen in example, in your answer.

Answer: Grammaticality

Sentence: Falsely reassured, investors sent 11.4 million to the Newport Beach company.

Question: What is the reason, investors sent 11.4 million to the Newport Beach company?

Answer: investors were falsely reassured

In order for the answer to sound more natural we prefer it if you try to answer as a full sentence. While it might be easier to just answer “falsely reassured”, it sounds more natural to answer “investors were falsely reassured”.

Similarly to the question span, try to stick to the “subject verb object” format in your answers.

Note, that in order to make the answer grammatical you are again allowed to introduce function words (no new content words!), such as we did with “were”.

Answer: Connective/Trigger

Sentence: I will do this unless someone complains.

Question: Unless what will I do this?

Answer: unless someone complains

Sentence: You could either go to sleep or play a bit longer on your phone.

Question: What is an alternative to you going to sleep?

Answer: or you could play a bit longer on your phone

Many (but not all!) sentences contain useful words that **indicate** what kind of a relation holds between two events, such as “because, or, after, while, despite” etc.

In cases where you can detect such a connective, try to include it in the beginning of your answer. In cases, where it doesn't sound natural to include it in the answer (see example 2), you should still include it in the beginning of the answer, but then proceed with the way you naturally would have answered the question.

Answer: Special Cases: Connective/Trigger

Sentence: Falsely reassured, investors sent 11.4 million to the Newport Beach company.

Question: What is the reason, investors sent 11.4 million to the Newport Beach company?

Answer: investors were falsely reassured

Sometimes there won't be any connectives to include in the answer. In those cases proceed as usual with answering the question. **Do not add a connective that isn't found in the sentence.**

Answer: Span

For the answer span try to keep the same principles in mind as for the question span:

- The **meaning** of the Question and Answer is **true** according to the sentence.
- The answer **does not include context that doesn't contribute to the meaning** of the Question and Answer.
- The answer **focuses around 1 main event/statement**, but might include other events if they are essential for the QA pair to make sense.

Answer: Span - Example

Sentence: The dam replaces an old dam, first built in 1911 but subsequently expanded several times.

Question: Before what was the old dam first built in 1911?

Answer: The old dam was subsequently **expanded** several times

Answer: The dam **replaces** an old dam.

When two events that are self-contained answer the same question, it is better to ask the same question twice, with two different answers.

Answer: Span - Example

Sentence: Mr. Van de Kamp is the one who collected the plans from the various radical environmental groups and cobbled them into a single unwieldy initiative to be placed on the ballot for election on Nov. 6, 1990.

Question: What is the reason Mr. Van de Kamp cobbled them into a single unwieldy initiative?

Answer: to be placed on the ballot for election on Nov. 6, 1990

Simply answering with “for election” would not have been enough. Writing the whole phrase fits the sense of the Question better.

Examples

Positive Examples: Multiple Questions

Sentence: Donald Trump withdrew his \$7.54 billion offer for American Air, citing the "recent change in market conditions."

Question: After what did Donald Trump withdraw his \$7.54 billion offer for American Air?

Answer: the recent change in market conditions

Question: What is the reason Donald Trump withdrew his \$7.54 billion offer for American Air?

Answer: the recent change in market conditions

Here you can ask two questions for the relation between "withdrew" and "change": a temporal one and a causal one. Also notice how for the first question, the copied word "withdrew" had to be changed to "withdraw", while for the second question it remained the same.

Negative Examples: Multiple Questions

Sentence: Their mission is to keep clients from fleeing the market, as individual investors did in droves after crash in October.

Question: What is similar to clients fleeing the market?

Answer: as individual investors did in droves after the crash in October

Question: What is similar to individual investors fleeing in droves after the crash in October?

Answer: clients fleeing the market

For relations that are symmetric, meaning that you could ask the same question start with either of the targets in the question, **you need to only ask 1 of the 2 possible questions! Feel free to chose yourself which of the two question options you prefer to ask.**

Negative Examples: Multiple Questions

Sentence: I ate dinner after coming home from work.

Question: After what did I eat dinner?

Answer: after I came home from work

Question: Before what did I come home from work?

Answer: I ate dinner

For relations that are reverse, meaning that you could ask the opposite question start, by reversing the targets in the Question and the Answer, **you need to only ask 1 of the 2 possible questions! Feel free to chose yourself which of the two question options you prefer to ask.**

Positive Examples: Different span for different question

Sentence: They wanted a court to block today's scheduled liftoff long enough for them to ask the World Court to order a permanent cancellation of the \$1.5 billion flight.

Question: Until when do they want a court to block today's scheduled liftoff?

Answer: long enough for them to ask the World Court to order a permanent cancellation of the \$1.5 billion flight

Question: What is the reason they wanted a court to block today's scheduled liftoff long enough?

Answer: for them to ask the world Court to order a permanent cancellation of the \$1.5 billion flight

Negative Examples: Too many modifications

Sentence: Blue Arrow of Britain plans to return to the name Manpower and take a big write-off.

Question: What is the result of Blue Arrow of Britain's plans to return to the name Manpower?

Answer: The company will take a big write-off.

In the question the verb 'plans' was changed to a noun. It is preferred to instead ask the question the following way: What is the result of Blue Arrow of Britain **planning** to return to the name Manpower?

In the answer a **content word was introduced that does not appear in the sentence**. Instead you should have answered this way: Blue Arrow of Britain will take a big write-off.

Negative Example: Wrong meaning

Sentence: Scotland's Police Investigations and Review Commissioner (PIRC) has criticised the national force for failing to stop a driver reported to them for being drunk behind the wheel.

Question: What is the reason Scotland's Police Investigations and Review Commissioner (PIRC) has criticised the national force for failing to stop a driver?

Answer: The driver was reported to them for being drunk behind the wheel.

Here the meaning of the Question and Answer does not correspond to the meaning conveyed in the sentence! They weren't **criticised** because the driver was **reported**, instead they were **criticised** for **failing** to stop the driver who was reported.

Negative Example: Too much context

Sentence: Under the agreement, Intel will invest \$3 million to acquire a 4% stake in Alliant, a maker of mini supercomputers for scientists and engineers.

Question: After what will Intel invest \$3 million to acquire a 4% stake in Alliant, a maker of mini supercomputers for scientists and engineers?

Answer: Under the agreement

In this negative example, the question span copied from the sentence is far **too long**. The information that Alliant is “a maker of minisupercomputers for scientists and engineers” is **extra context** that **does not belong to the meaning of this specific QA pair**, same for the purpose “to acquire”, **which is an event in itself!**

The correct question would have been “After what will Intel invest 3\$ million?”

Don't be afraid to say “No Question applicable”

Sometimes you might receive targets that really **don't fit any of the questions** we want you to ask. In those cases, instead of trying to invent a question that doesn't really fit the meaning of the sentence and the target(s), click on “no question applicable”. In the example below for example there is no obvious question to ask for “according”.

Sentence: **According** to the Team Manager Nick Sullivan, Smith took a while to settle in before taking her first shot.

Question: No Question Applicable

Process

Training and Qualification

We will release two training rounds:

1. **For the first round** you should carefully read these guidelines here and the guidelines in the interface. Then, you'll be given 31 HITs which you should complete according to the guidelines. We will then assess your work and send you back feedback and suggestions on how to improve your annotations.
2. **For the second round** you should re-read these guidelines and the feedback we gave you on the second round. You'll then be given another 30 HITs. If we see that you follow the guidelines and that you did well on those 30 HITs of the second round, you'll be granted a qualification to start working on the production batches (which will be large batches of thousands of HITs in total).

Compensation and Bonuses

- Each HIT will be rewarded with ¢18
- Additionally, you will get a ¢3 bonus for an additional question after the 1st question and ¢4 for every question after the 2nd.
- A \$4 bonus for reading the guidelines, our feedbacks and completing both training rounds in a timely manner.

To sum up

If you have any questions please contact us at the following address:

[anonymized](#)