Scheherazade Tutorial for Annotating Blog Stories

Part 1: Introduction to Story Intention Graph (SIG) and Scheherazade

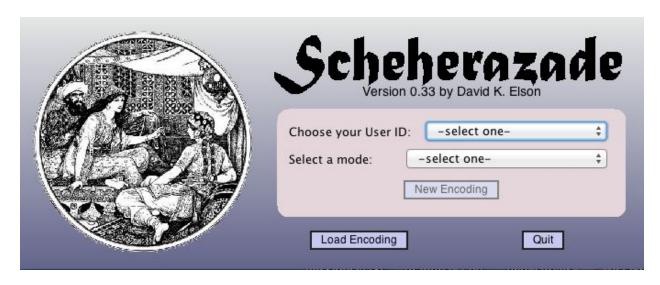
Scheherazade is a software tool and platform for symbolically encoding narratives using the SIG representation.

"Symbolically encoding" means taking two kinds of input, a source text and a controlled vocabulary of structured data (a kind of dictionary of story elements), and recreating the story as well as possible using only the story elements in the controlled vocabulary. It's kind of like translation, except you are translating into a machine-readable language rather than another natural language.

The controlled vocabulary that Scheherazade uses (SIG) assumes that stories are made of:

- **nouns**, which can be found in an electronic noun dictionary
- **verbs** and the selectional restrictions of their semantic roles, which can be found in an electronic verb frame dictionary
- adjectives, adverbs and other **modifiers**
- a notion of the underlying **timeline of a story**, including states in time, and references to hypothetical actions that don't actually take place
- a strong sense of **agency** among characters in the story: that they have inner worlds and desires
- **goals** that those characters have
- plans that those characters devise to reach those goals
- **attempts** by those characters to reach their goals
- **outcomes** of those attempts and the resulting affectual impacts

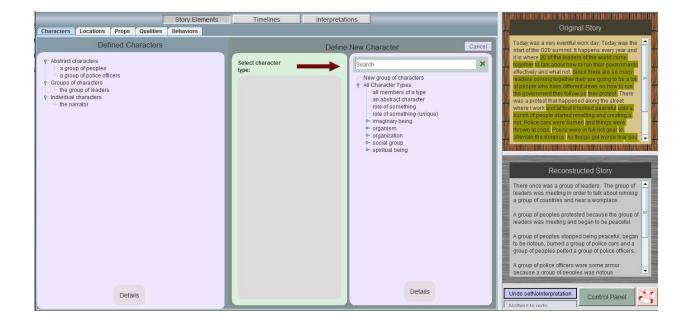
Part 2: Overview of the Editor



For "Select a mode", you can select to either create a:

- **Story From File:** This will prompt you to select a plain text file.
 - This is the option we choose for annotating blog stories.
- Original Story: This will start Scheherazade without a plain-text story for annotating.

Click "New Encoding" to start the editor. If you have already started and saved a story, click "Load Encoding" instead to continue work on your encoding.



Part 3.0: Understand the basic idea of the story

Story to annotate: Today was a very eventful work day. Today was the start of the G20 summit. It happens every year and it is where 20 of the leaders of the world come together to talk about how to run their governments effectively and what not. Since there are so many leaders coming together their are going to be a lot of people who have different views on how to run the government they follow so they protest. There was a protest that happened along the street where I work and at first it looked peaceful until a bunch of people started rebelling and creating a riot. Police cars were burned and things were thrown at cops. Police were in full riot gear to alleviate the violence. As things got worse tear gas and bean bag bullets were fired at the rioters while they smash windows of stores. And this all happened right in front of my store which was kind of scary but it was kind of interesting since I've never seen a riot before.

These stories are rich in language and description, and have many things that aren't relevant to the annotation of the story events. Before you begin annotating with the tool, look at the story and determine the characters and events that are crucial to the plot. This story is fairly straightforward and most of the blog is actually describing the events that take place.

Characters:

- Narrator
- Leaders invited to G20 summit
- Protestors
- Police

Props

- Police cars
- Riot gear
- Tear gas
- Bean bag bullets
- Store windows

Locations

- Street
- Store narrator works in

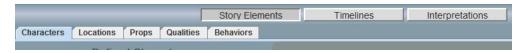
Actions:

- Leaders are gathering for G20 summit
- People begin to peacefully protest
- Protesting turns into a riot
- Police cars are burned
- Police come to alleviate violence
- Tear gas and bean bag bullets are fired at protestors

Part 3: Declaring Nouns

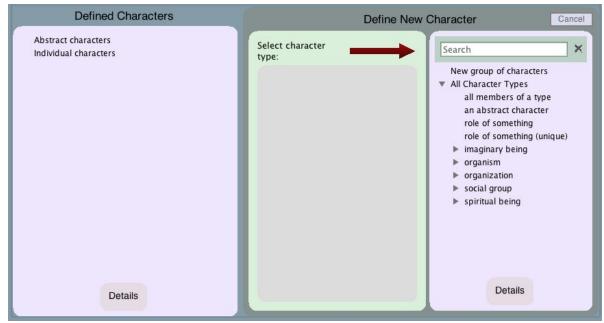
(Further reading: https://sites.google.com/site/scheherazadetutorial/declaring-nouns)

In 'Story Elements', select 'Characters'



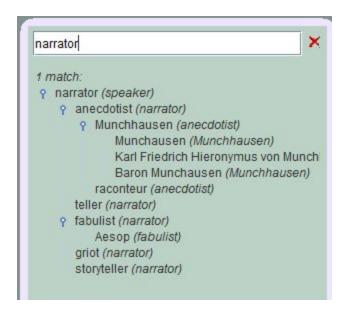
This is divided into two smaller panels:

- **Defined Characters** is where you'll see a list of characters already declared in your story.
 - O Individual characters include discrete individuals like the Crow in our story, as in "I found <u>a crow</u> at the pet store."
 - O Abstract characters are classes that you want to refer to in your story, as in "I think I'll go adopt <u>a crow</u>." The phrasing looks remarkably similar to that of an individual character, but if you squint one eye and look very close, it carries a different meaning in our semantic encoding.
- Define New Character is a worksheet where you can do just that. This is your first experience with a form. Forms are used throughout Scheherazade to set up various kinds of instances. In this case, we are going to define an instance of a character type. Every time we set up a form, we'll see two smaller panels:
 - On the right will be a hierarchy of **types** -- in this case types of characters -- and a search box.
 - On the left will be a place to configure the instance and finalize it. We'll see how in a moment.



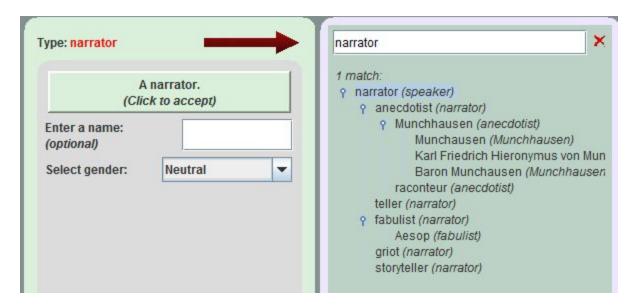
All the blog stories are told in a 1st person voice. We define a **narrator character** to be that voice.

- Note that there is a label "Select character type" with a red arrow pointing to the search box.
- Type in "narrator". You'll see results like these:



NOTE: Search results will appear as you type, kind of like Google Instant. Sometimes the search function can be a little buggy, in that if you type too fast, it will lose track and not show results consistent with your query. If this happens, just delete and retype the last character in your query. The search function will retrigger and display the correct results.

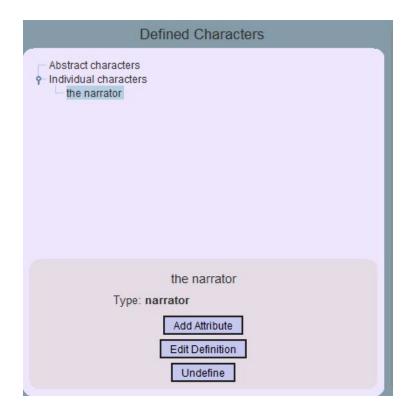
- You'll see all search results that contain the string "narrator", arranged hierarchically where appropriate. For each choice, the "genus" (category) is shown in italics. You can click any of these results, even if it has sub-results. So let's click on the closest match, "narrator (speaker)".
- Here, as in any form situation, when you select a type, the sub-panel on the left will show the slots that need filling for that type, as well as an accept button on top that will allow you to complete the definition.



In this case there are two slots that need filling:

- Name, which is optional (you can leave it blank)
- Gender, which is used to determine which pronoun to use in retelling the story (he, she or it)
- The accept button reads out the result of the form in natural language for you to approve. In this case, it's asking if you are happy with defining "A narrator."
- If you want, you can select a different type from the search results panel. The slot-filling panel will update with the new slots for the new type. But in this case, let's be happy with a narrator.
- The narrator in the story is of an unknown gender. Select a gender if it is specified, and if it is not, select "Female" or "Male" from the dropdown menu.
- Click the accept button.

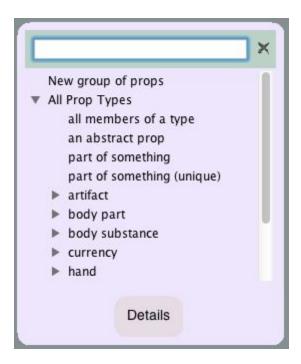
When you hit the accept button, the "Define New Character" panel resets to its blank state, but the "Defined Characters" panel now includes one item underneath "Individual characters".



Now the defined narrator is here. The informational panel at the bottom shows you a few things:

- The type (in this case, "narrator" was the type you selected from the search results panel)
- An Add Attribute button, which we'll leave until later
- An Edit Definition button, which will cause the "Define New Character" form panel to trigger into an edit mode where you can change anything about the defined crow, such as its type, its gender, or its name, and then approve your changes withe the accept button
- An Undefine button to delete this character definition, removing the crow from your encoding

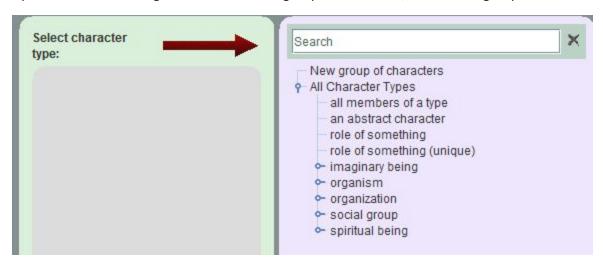
There are several special noun types that are available for all noun classes (characters, locations and so on):



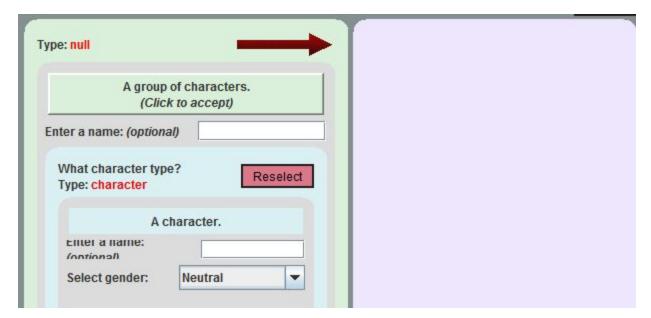
- 1. New group of nouns (in this case, "New group of props") is a special type for defining groups. You can define an anonymous group ("a group of foxes") or assign particular members to the group.
- 2. All members of a type. Use this to refer to all individuals of a type collectively ("every fox").
- 3. An abstract type. Remember, above, where the list of defined characters included both abstract characters and individual characters, which looked remarkably similar even after you squinted? Use this special type to define an abstract noun, such as an abstract prop ("he decided to buy <u>a computer</u>").
- 4. Part of something. Use this to refer to a <u>meronymous</u> relationship, in which the whole can have more than one instance of the part ("a branch of a tree").
- 5. Part of something, unique. Use this to refer to a meronymous relationship in which the whole can only have one instance of the part. (We want "the beak of the crow", not "a beak of the crow").

In all five of these cases, when we create the noun, we have to do more work than just giving a name and a gender

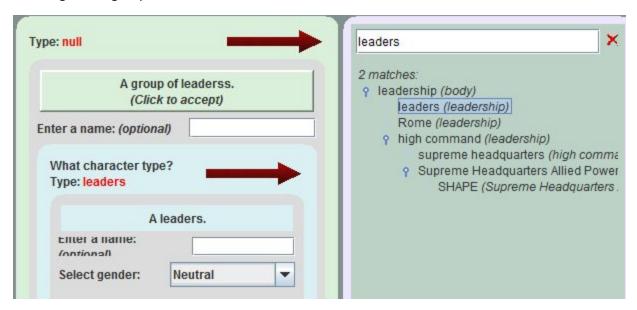
Groups are common in blog stories. To define a group of characters, select 'New group of characters'



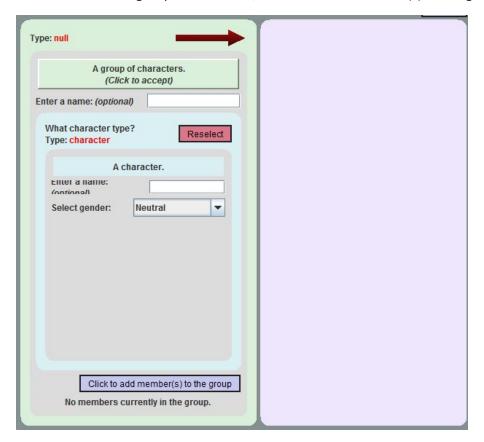
Reselect 'what character type'



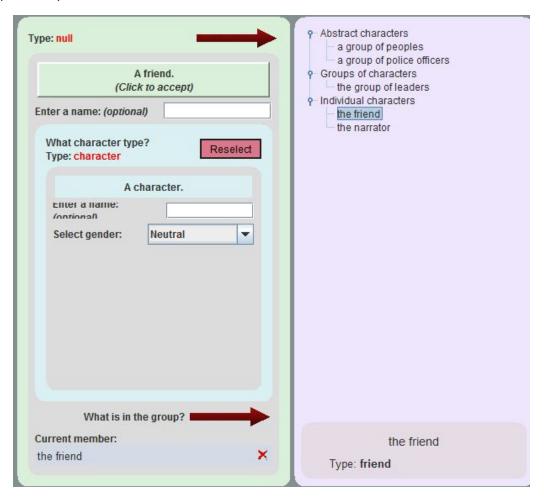
This is a generic 'group of leaders'



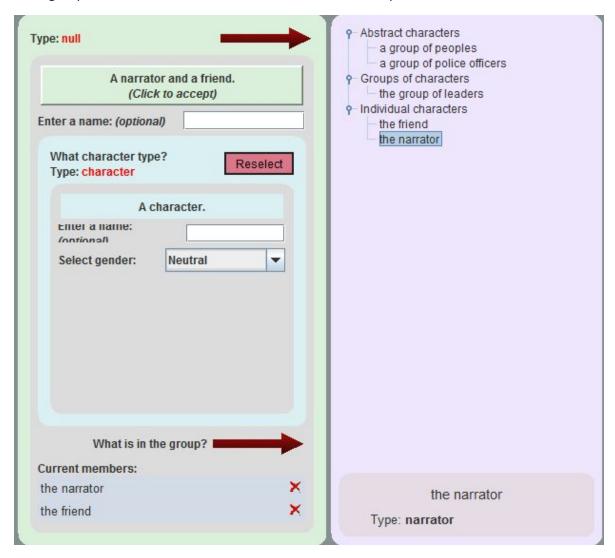
If the group is more specific and contains previously defined character, we can add members to the group. At the bottom of the 'new group of character', do 'click to add member(s) to the group'



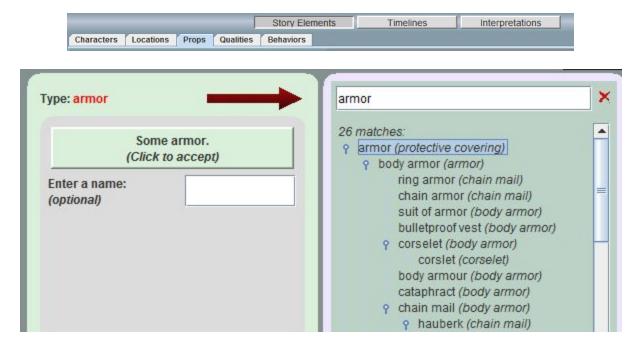
Select previously defined characters



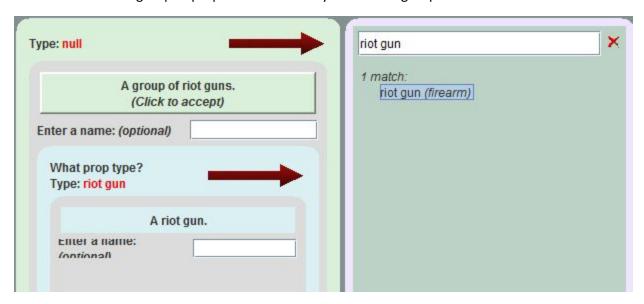
And the group will be referred to as the members who make it up.



In addition to characters, we can define props. Select the 'props' tab



We can also make a group of props in the same way we make a group of characters



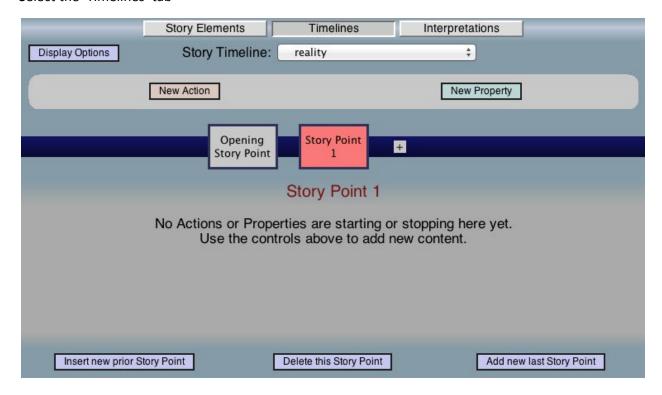
Note there are other nouns to define including locations, qualities, and behaviors.

Part 4: Declaring Actions and Properties

(Further reading:

https://sites.google.com/site/scheherazadetutorial/part-04-declaring-actions-and-properties)

Select the 'Timelines' tab



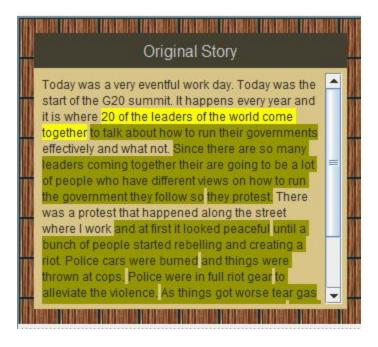
Here's what's going on:

The Timeline. The big blue bar represents the enduring march of time, linear and infinite (at least in the forward, or future-facing direction). We can add an arbitrary number of Story Points that represent instants in time (states). We add as many as we need. A Story Point represents an arbitrarily small moment of time. Nothing actually takes place at a Story Point. Instead, everything in the story takes place in the intervals between Story Points. Actions can overlap; for example, I may be walking down the street in the interval between Story Points "2" and "6", and thinking about my purpose in life between Story Points 1 and 4 (partially overlapping). Note, too, that there is no meaning to Story Point numbers except that they are sequential, like slide numbers in PowerPoint or another presentation tool. They don't represent days, or minutes, or hours, and there is no way right now in Scheherazade to define the lengths of real-world time that transpire in the interval between Story Points. All that really matters in our representation of a timeline is the interval logic: what actions take place over what spans, and how they relate to each other.

- The buttons along the bottom let you insert or delete Story Points. Think of this as like inserting or deleting slides in the presentation tool. In that case, you have a linear sequence of slides, and you can insert new slides in the middle, and all the slides afterward are renumbered. Here, if you add a new Story Point prior to Story Point 1, you'll make a new Story Point 1, and what was Story Point 1 becomes Story Point 2. Alternatively, you can click the + button after Story Point 1, and a new Story Point 2 will be created and appended to the timeline. You can create as many Story Points as you like. Try it out; it's kind of satisfying to make a bunch and then delete them. Sometimes during the more tedious nights of working on my dissertation, I would pass the time by creating and deleting hundreds of Story Points, over and over, watching time dilate and compress, seeing the ether of existence proliferate and wither, acting like the sand-castle builder and the sum of the ocean's renewing waves alike....

 Just kidding, that would have been rather desperate of me.
- The "Display Options" button in the top left can control some cosmetic features that won't make sense until we make a few actions.
- The story timeline has a name: Reality. More to the point, this is the Reality timeline. There are other lines of time, you ask? Well, there are if you construct them. For example, we can model imaginary actions, like fears and hopes, as alternate timelines, and then attach them to Reality. For more details, see the full tutorial (https://sites.google.com/site/scheherazadetutorial/part-04-declaring-actions-and-properties)
- Finally, two buttons allow you to create new content: New Action and New Property.
 - O Action corresponds to a <u>dynamic verb</u>, i.e., something progressive that looks different throughout the interval (like taking a seat)
 - O Property corresponds to a stative verb, i.e., a state of being that is static throughout the interval (like sitting)

Let's declaring an action for the phrase "20 of the leaders of the world come together". First highlight the appropriate text span in the story

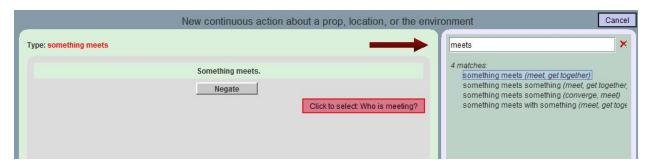


Click 'New Action'. The Timelines screen is taken over by this form-filling panel. It should feel pretty familiar by now; it uses the same underlying code as the noun-defining panel we used earlier. The difference is that instead of defining a new noun, we are defining a new Property and putting it on the timeline. A new Property or Action will be attached to the Story Point that was selected when you first clicked New Action or New Property, in this case Story Point 1.

There is a difference compared to earlier, though, and that is the "Negate" button. Try it out: Our placeholder is now negated.



We want to annotate the idea of "coming together" or "meeting". Let's start by typing "meets". A list of possible verb senses and arguments are shown. Select the one most appropriate to your text span. The 'somethings' are slots that we will fill in.



There is a red prompt button: "Click to select: Who is meeting?" Clicking that presents a picker that should be familiar. We can select all the various things that can begin to sit in Story Point 1 of our encoded story:

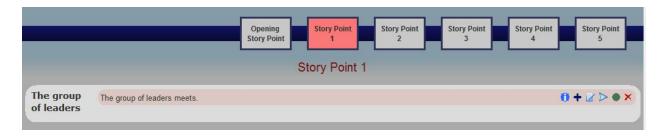


Select "the group of leaders" and the action is ready to go.

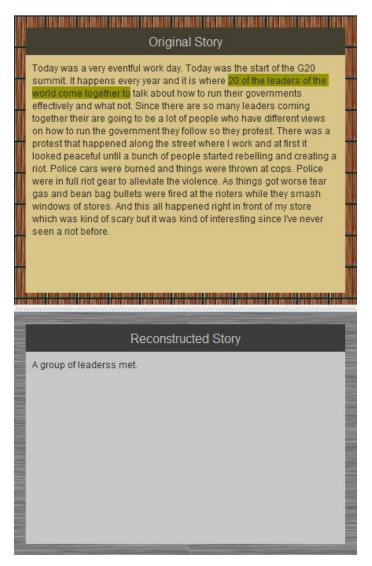


Note how the answer to the prompt "Who is meeting?" is now labeled where the prompt button used to be: "the group of leaders".

Now click the accept button on the action panel. You've now created the action <u>and</u> associated it with the span of text above. You should see the span go from red to tan, and you should see your new action appear below the timeline:



As you click around the Timeline, the actions and properties that start or end at that Story Point appear in the space below the timeline. Here, because the act of sitting starts at Story Point 1, we see "The crow begins to be sitting." We'll talk about those controls on the right momentarily. Here is how the text panels look now:



The **Original Story** panel now has the sentence we have already modeled highlighted in tan; this gives you a sense of how much of the story you have already encoded. The **Reconstructed Story** panel has gone from empty to a reading of our entire story as we've modeled it symbolically. It's not much so far, just the one predicate we've set up. But eventually it will be a machine reading of a large semantic network that represents our reading of the fable.

Now, about those five control buttons to the right of the sentence in the Story Point area:

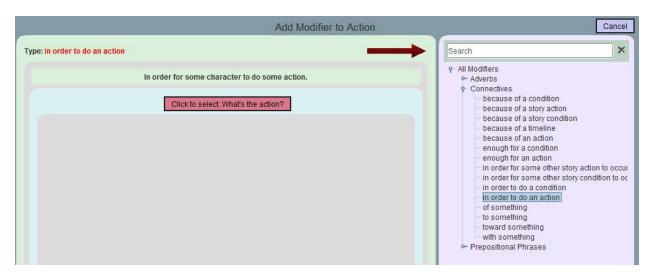


- The i button is for information. It tells you a little bit of information about how you set up that sentence. Here, it tells us that "The crow begins to be sitting" is "a continuous property that begins here and doesn't stop", i.e., the crow starts to sit in Story Point 1, and never stops sitting. (Thrilling story! Better than *Cats*!)
- The + button lets us add a **modifier**. We'll do that momentarily to talk about the crow in the beak.
- The paper-and-pencil icon is for **edit**. If you click it, the form creation panel re-appears with all the slots pre-filled as we filled them out earlier. If you want to change the "crow sit" property, you can hit edit, make your changes, then click the accept button again. Note that you can also add a new Original Story span or change the span using the edit button. There isn't anything telling you this in the interface, but if you select a new span while editing an action or property, you'll create a new association when you click the accept button again.
- The triangle button lets you **move** an action or property to a different Story Point.
- The **X** lets you **delete** an action or property.

Let's first tackle that modifier. Hit the + button and you'll see a new form-filling panel, this one for **modifier** predicates. Here, you can specify that the group of leaders meet in order to talk about running the countries. Instead of searching for something, just click the triangle next to the "Connectives":

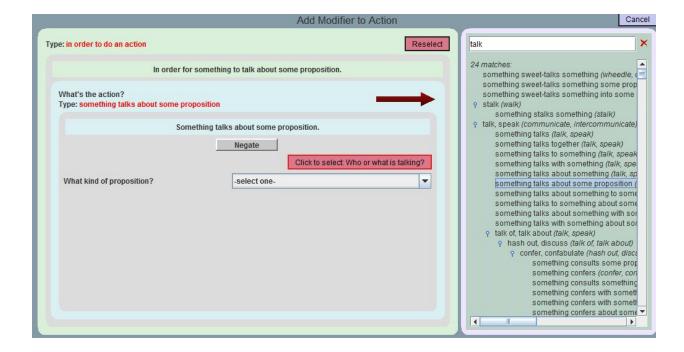


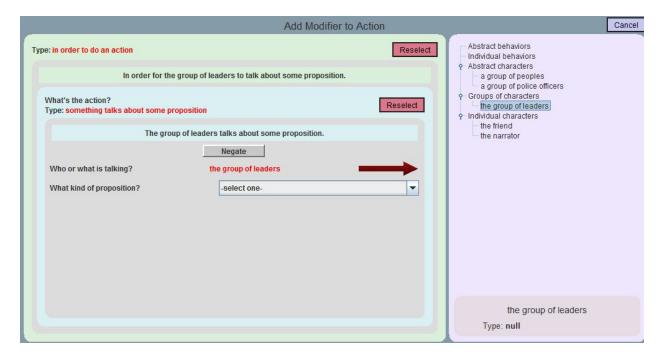
Click on "in order to do an action"

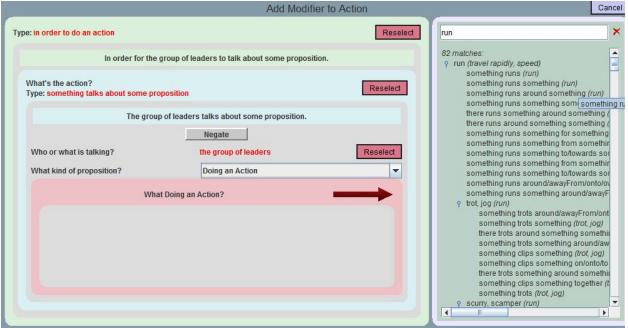


Fill out the form:

- Type "talk" as the action (don't forget to click on the red prompt button first)
- Select "the group of leaders" for the subject (first click on the red prompt button)
- Highlight "talk about how to run their governments" in the Original Story
- Select "do an action" under "What kind of proposition?"

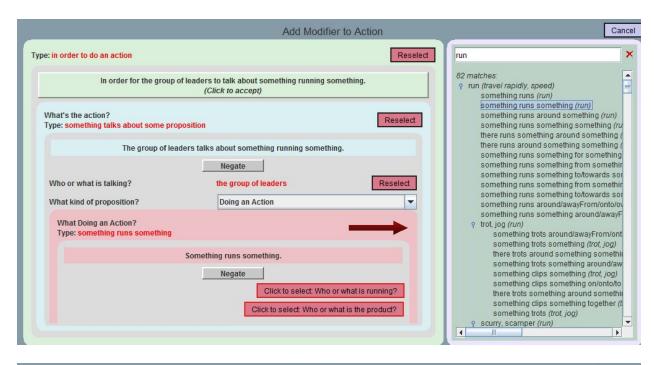


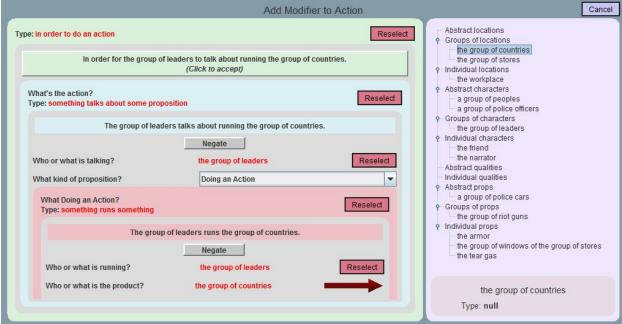




This is an embedded action. Again, we fill out the form:

- Type "run" as the action (don't forget to click on the red prompt button first)
- Select "the group of leaders" for the subject (first click on the red prompt button)
- Select "the group of countries" for the product

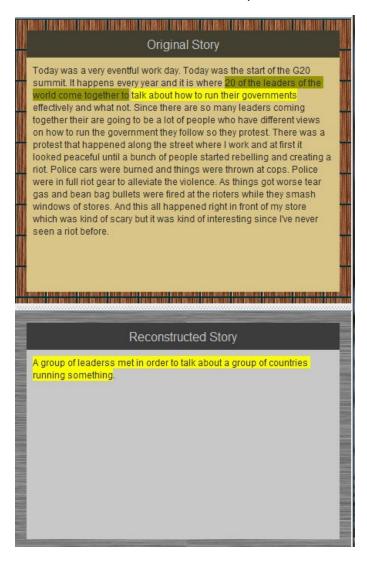




You should now see a modifier dutifully placed in the Story Point area:



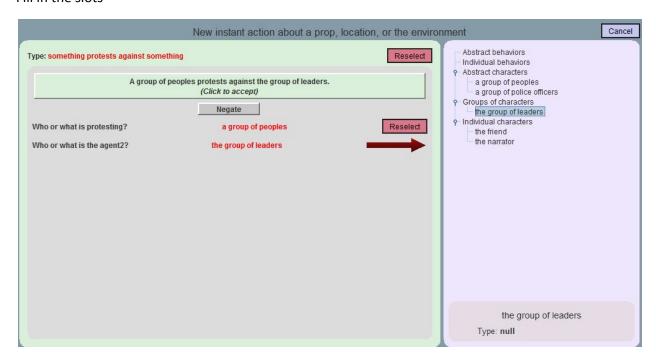
And an inclusion of the tree branch in the Reconstructed Story:



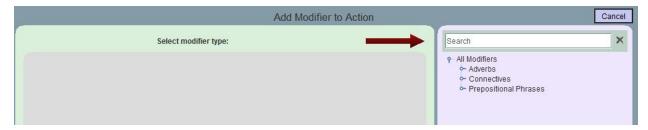
Let's do another example. "The people protest against the leaders"



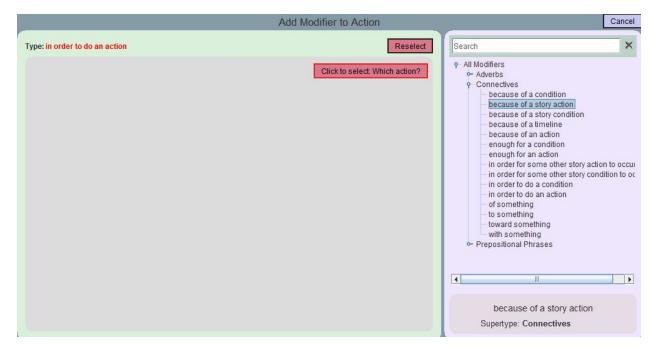
Fill in the slots



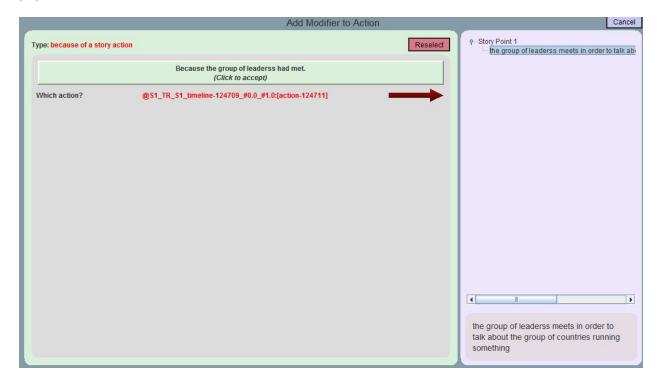
Click to accept and add a modifier to be able to say **The people protest against the leaders <u>because</u> <u>the leaders are meeting</u>**



Again select connectives and "Because of a previous story action"



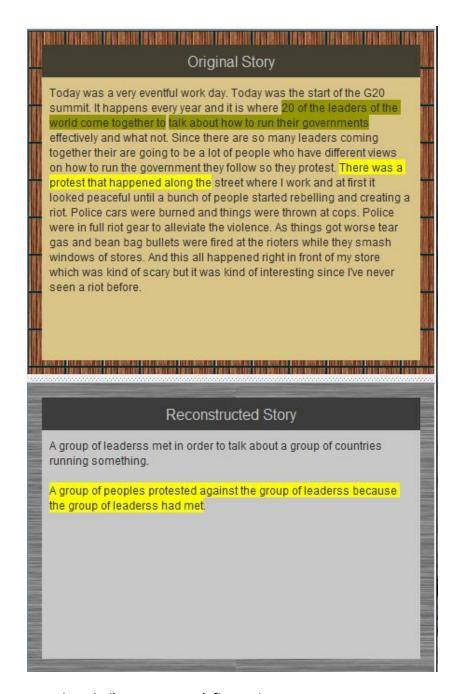
"Click to select: Which action?" gives us the story points we have so far defined. We select the first one.



And "click to accept" and have our story point.



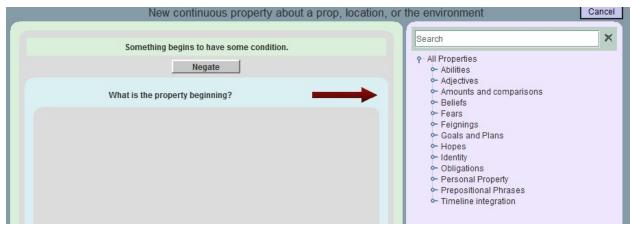
Keep an eye on the reconstructed story. If you click on a story point, the corresponding text in this box will be highlighted. This text is here so you can see how good your annotations are and if you are annotating what you intended to say.



Let's define a property in a similar way as we define actions.

A note on terminology: You'll usually see "property" but sometimes you will see "condition" for obscure reasons. They are interchangeable. For example, the accept button banner here reads "Something begins to have some condition", i.e., there is some Property that is as yet undefined. When we define the Property, the banner will pop up into a clickable accept button, just like in the case of nouns earlier.

Create a new Property



Notice that the question being asked is: "What is the property <u>beginning</u>?" (Emphasis mine.) Even though Properties are stative, they still have beginning points and end points. The phrasing of this question is just to emphasize that we are actually modeling the onset of a Property, which you can think of as a special kind of Action (in that it is state-changing). At Story Point 0, the Property does not hold; by Story Point 1, it has commenced.

Now, what is the stative Property we are trying to model? That the people are protesting peacefully. Try typing "peaceful" in the Property search panel.

Now the accept button banner can be a bit more specific and say: "Something begins to be peaceful." (Awkwardly phrased, but technically correct given how we are modeling this.)

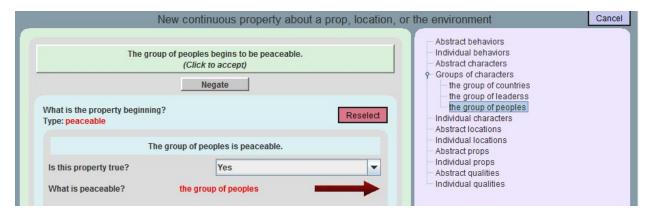
Here we have a nested form in a blue box that represents the Property that is beginning. It asks if "the property is true". But wait, you say, don't we implicitly ask that with the "negate" button? Well, the "negate" button refers to the <u>onset action</u>, where the "property true?" dropdown refers instead to the property itself. In other words, four combinations are possible:

- the group of people is peaceful
- the group of people begin to not be peaceful
- the group of people aren't peaceful
- the group of people doesn't begin to not be be peaceful

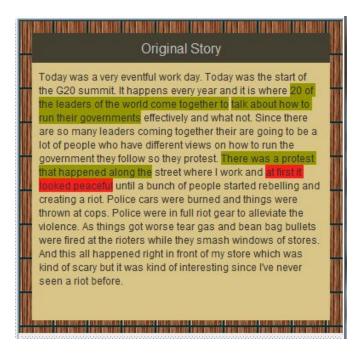
At first blush, this seems rather redundant semantically, but it can be useful to have this distinction in certain cases that are less clear-cut than being peaceful. Consider the difference between "the trader didn't become rich" and "the trader became not rich". The semantics are quite different. For the same reason, the "is the property true?" dropdown also allows for degrees: one can model that the trader became extremely rich, somewhat rich, very not rich, and so on.

In this case, let's leave the property true and the onset non-negated.

There is one last red prompt button: "Click to select: Who are peaceful?" Clicking that presents a picker that should be familiar. We can select all the various things that can begin to be peaceful in our encoded story:



Select "first it looked peaceful" from the Original Story



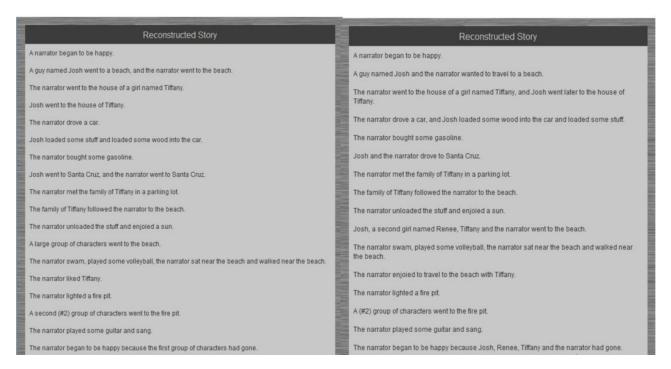
Now click the accept button on the property panel.



An Example Annotation

The protest story annotation provided to you (a034_protest.vgl) is an example of very good annotation.

The reconstructed story on the left (1123014_bonfire.vgl) is an example of a basic annotation that could use some improvements. The reconstructed story on the right (1123014_bonfire-edit.vgl) makes some of these changes.



Original story for 1123014_bonfire:

"Today was so much fun. I and Josh planned to have a beach day/ bonfire at Rio Del Mar. I met up at Tiffany's house and I started from there. Josh met me there too. I drove my Mustang and Josh loaded up all of his stuff and put some wood in the back.

I started out driving but I needed to get gas so they started driving and I filled up my tank. Then I and Josh cruised down to Santa Cruz. I met up with Tiff's family in the parking lot of the Walgreens and then they followed me to the beach. I had to haul all the stuff down to our spot and then set up. It was a really nice day and I got to be out in the sun. I actually had the biggest group there on the beach. Renee and Jojo came. Char and Julie, Lucas, Travis and just about everyone else. Jacob Esquibel came and Jake Aguilar.

I went into the water and body surfed. I played volleyball and sat by the beach. I even took a long walk along the beach and it was really great to be around Tiffany so much! I started the fire as the sun was going down and had it going pretty good. Shannon, Carlos, and Rodney came too and there was a

bunch of people from Carmel and Monterey. It was nice. Then out came the guitars and I played and sang. It was a great night and there were a lot of people there. It was so much fun."

Take a look at the annotation for **1123014_bonfire.vgl** and see the comments below:

The annotator was able to capture all the timeline points and introduce most major characters, props, locations. Here are my suggestions for the bonfire story:

- Use a wider vocabulary, especially on the verbs. Instead of 'went to the beach' --> 'traveled to the beach'; 'went to Santa Cruz' --> 'drove to Santa Cruz'.
- Create individual characters and add them as members of a group. 'A group of characters' -->
 'Josh and the narrator'
- Try to join some story points together. 'The narrator drove to Santa Cruz. Josh drove to Santa Cruz.' --> 'Josh and the narrator drove to Santa Cruz'
- Be more specific. 'The narrator liked Tiffany' --> 'The narrator enjoyed traveling to the beach with Tiffany' (Realized as 'The narrator enjoied to travel to the beach with Tiffany')

Some of these changes will produce strange realizations in the reconstructed story (the above 'enjoied' and introducing characters into groups, which usually says "a girl named Renee"), but we have ways of working around them.

Take a look at the changes in **1123014_bonfire-edit.vgl** (but note that I didn't change everything, nor is the above list inclusive of all the things I changed).

Then take another look at **11230214_camping.vgl** (reconstructed story below) and note ways to improve the annotation. Groups appear very often in these blog stories, so it's important to properly introduce characters and give them names. Some concepts are hard to annotate (e.g. The narrator enjoyed hanging out with Tiffany) but we have to try to better explain concepts and use wider vocabulary.

	Reconstructed Story
A group of charac	ters travelled to a camp.
The group of cha	racters had some food.
The group of cha	racters drank some alcoholic drink and began to be sick.
The group of cha	racters ate some watermelon.
A narrator ate the	slice of the watermelon.
The narrator swa	m around some rope.
The narrator met	a crazy group of characters.
The group of cha	racters shot some water with a firework.
The group of cha	racters didn't hit the narrator and lighted the firework.
The narrator left t	he water and walked toward the group of characters.
The firework stru	ck the narrator.
The narrator gave	the group of characters a number and left.
The group of cha	racters began to be drunk and went towards a park.
The group of cha	racters fought with a guy, and the guy scratched a car.
The narrator play	ed a playing card.
A second (#2) gu	y began to be drunk.

Original story for 11230214_camping:

"Jay, Matt, and I went camping this weekend. I was well provisioned, Matt had everything you could think of for the trip: a huge tent, hotdogs, marshmallows, coolers, ever clear, vodka, a watermelon plugged with ever clear, among other things. The ever clear ended up being too much for us, and Jay and I were both holding our stomachs before long. Watermelon did little to soften the effect, but it was amusing watching them attempt to keep a straight face as they bit into it. In all honesty, I was probably more amusing to watch, but I managed to force down a piece. Unfortunately, I had gotten a piece particularly close to the plug.

I took a long swim over to the rope swing where I met up with some crazy guys shooting fireworks into the water. After nearly hitting me, they apologized, but continued lighting the fireworks, anyway. I eventually climbed out of the water to join them and the craziness that ensued left me all with at least one close encounter with a bottle rocket (or in some cases, straight on hits) and exchanging phone numbers before the day was over. I think I probably would have hung out longer right? But they got really drunk and left for a park in order to pick a fight with some guy that had keyed one of their cars. So I came back that night and played a few rounds of cards. Matt got completely trashed He's one of those that will repeatedly tell you he's not drunk, over and over again, and that's when you know he's pretty far gone!"

Some final tips:

- → Start annotation from **Timeline layer**: Identify key properties and actions of the story and add Story Elements (characters, props, locations, etc.) along the way as needed. This will help you focus on important parts of the story and annotate faster.
- → Define a new **Story Point** in the Timeline only when you need one (see 2 examples of bonfire story above in previous section).
- → If you can't find the exact terms of the story in the Scheherazade dictionaries, try to choose an appropriate substitute that conveys the same concept. Here is some examples:
 - ◆ Expletive "it": It was hard to ... → The situation was hard to ...
 - ◆ Pronoun "we": We decided to ... → A group of friends decided to ... (choose an appropriate group of characters based on the context of the story)
 - ◆ Use "groups" for plurals: five trees → a group of trees
- → Blog stories might contain some descriptive parts that are not easy to annotate and interpret using SIG representation. These descriptions mostly do not pertain to the key aspects of the story and in that case you can ignore them. For example, see the red highlighted text in the story below:

