



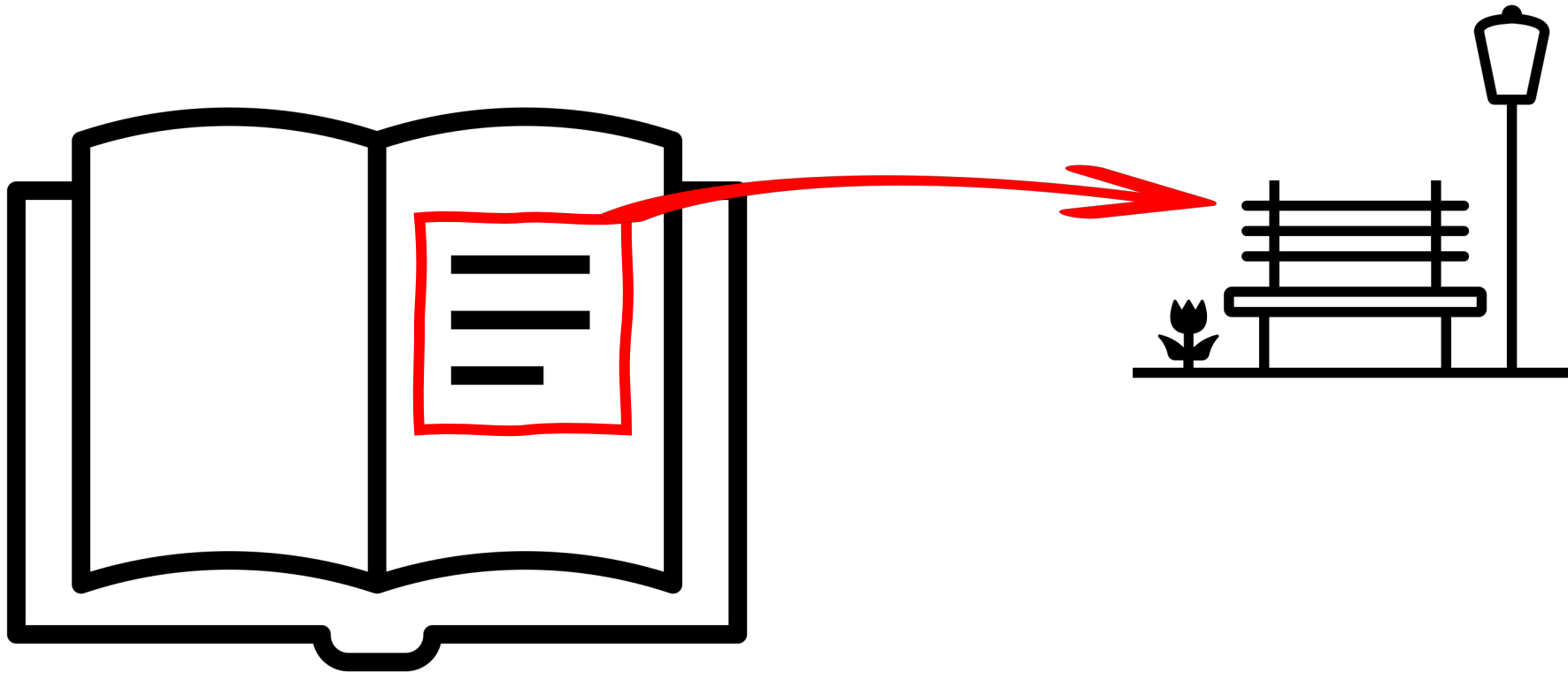
The  
University  
Of  
Sheffield.

# A Pilot Study on Annotating Scenes in Narrative Text using SceneML

Tarfah Alrashid and Robert Gaizauskas  
Department of Computer Science  
The University of Sheffield

# What if ?

We could automatically annotate scenes in narratives?



## Why do automatic scene segmentation?

To support:

- narratives studies
- automatic text illustration/story picturing
- aligning text and movies
- automatic image description
- automatic narrative generation



# Motivation

# Outline

- SceneML
  - There is a previous paper about sceneML, and this paper is about the pilot study and will only briefly explain what sceneML is
- Pilot study setup
- Annotation Example
- Results
- Discussion
- Conclusion and Future work





# SceneML

---

- A **scene** is a unit of a story in which the elements: **time**, **location**, and main **characters** are constant
- Any **change** in these elements indicates a change of **scene**



# SceneML Elements

---

## Entities:

- **scenes** – the main element in SceneML – abstract discourse elements
- **SDSs** (scene description segments) – the actual strings of text that comprise a scene – one or more per scene
- **characters**
- **times**
- **locations**



# SceneML Elements

## Relations

- Scene-scene narrative progression links:
  - **Sequence:** a scene change happens because of a change in location or characters, but the action follows directly on from the preceding scene
  - **Analepsis:** when there is a flashback in the scene, e.g. memory of the past
  - **Prolepsis (or flashforward):** when we are taken forward in time
  - **Concurrence:** when the transition happen because there is another **thread** of the story happening **at the same time** so the transition take us to different characters and different places at the same time
- Other relational links
  - character – scene
  - time – scene
  - location - scene



# Example

John and Tommy are walking down West Street heading to the supermarket

...

They remembered a year ago when they traveled together to Spain and had a lot of fun

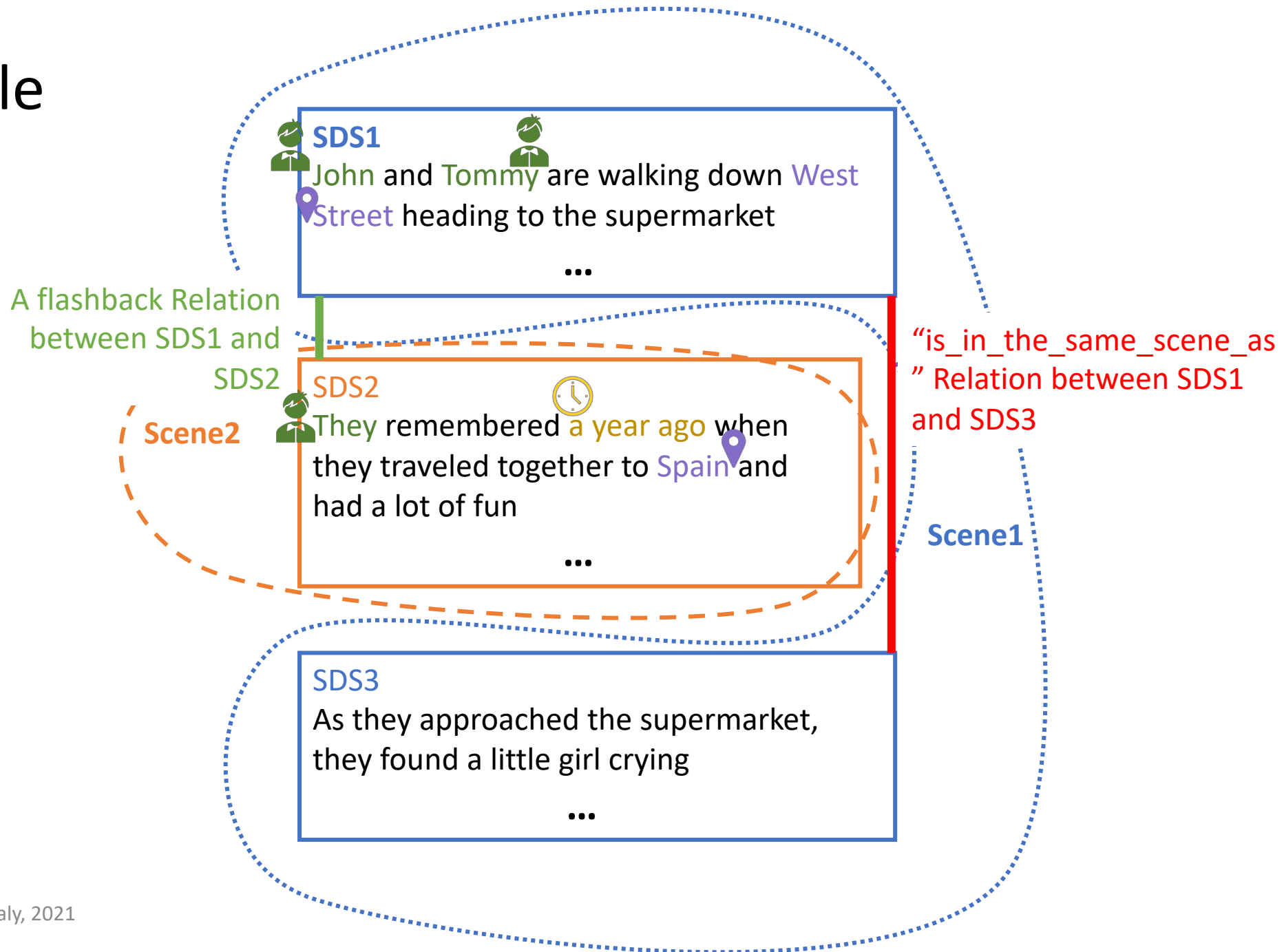
...

As they approached the supermarket, they found a little girl crying

...



# Example



# Pilot Study : Aims

- to **assess the viability** of the scene segmentation task and the **adequacy** of our guidelines.
- to determine **how** our annotation guidelines and processes should be **improved**
- to determine whether or not there are any underlying conceptual **problems** with our approach.



# Pilot Study Setup: Task

- **Two** chapters of children story “Bunnies From the Future”.
  - reading ages 10-13
  - chapter 3: 124 sentences long (2756 words)
  - chapter 4: 65 sentences (1775 words)
- **Four** annotators (postgraduate students) were asked to:
  - annotate SDSs, character, time, and location entities found in the given text.
  - follow the guidelines based on sceneML .
  - use **the Brat** annotation tool.

# Pilot Study Setup: Limitations

- The *scene* abstract discourse element is **not** annotated
  - Scenes are zero-span annotations and Brat does not support these
  - The problem can be circumvented by **linking all SDSs** in one scene together with a “**same-scene-as**” relational link
- **NO** relations annotated
  - for this first annotation exercise we were **primarily concerned** with determining whether or not annotators could accurately **identify** and agree on **SDS boundaries**
  - given that in our chosen texts virtually **all scenes** consist of a **single SDS**, the need to link multiple same-scene SDSs is low

# Annotation Example

	<div>CharacterTime</div> <div>Scene</div> <div>1 "Wakey wakey, Turtle. Time to rise and shine."</div>
	<div>Character</div> <div>Scene</div> <div>2 Skip's voice brought me back to the land of the living.</div>
	<div>Scene</div> <div>Time</div> <div>3 "Let's go and have breakfast, and I'll brief you on today's mission."</div>
	<div>Scene</div> <div>Location</div> <div>4 He unhooked me from the bunk,</div>
	<div>Character</div> <div>Character</div> <div>Location</div> <div>and I followed him down the corridor to the canteen. I lost control a couple of times on the way, and by the time we got there, I had a bump on the head, a scrape on my hand and</div>
	<div>Character</div> <div>Scene</div> <div>Location</div> <div>WingCo was just finishing his breakfast when we arrived in the canteen.</div>
	<div>Scene</div> <div>Character</div> <div>5 "How's he doing," he asked Skip.</div>
	<div>Scene</div> <div>6 "He got here all on his own." replied Skip, "A bit slow and a couple of bumps, but he made it under his own steam."</div>
	<div>Character</div> <div>Scene</div> <div>7 "Well done, Turtle," said WingCo, patting me on the shoulder, "Without your hard work yesterday, I swear you'd have had at least a broken arm – if you'd have made it at all."</div>
	<div>Character</div> <div>Character</div> <div>Location</div> <div>8 Skip and I collected our TCS from a dispensing machine and went to sit at one of the few tables in the canteen – a tight squeeze for me as the chairs were bunny sized. I noticed</div>

# Results

## Inter-annotator agreement results.

	Ann2				Ann3				Ann4			
	Ch3		Ch4		Ch3		Ch4		Ch3		Ch4	
	K1	K2	K1	K2	K1	K2	K1	K2	K1	K2	K1	K2
Ann1	0.60	0.54	0.33	0.27	0.15	0.27	0.20	0.30	0.23	0.19	0.10	0.21
Ann2					0.27	0.20	0.42	0.39	0.19	0.24	0.35	0.30
Ann3									<b>0.72</b>	0.33	<b>0.95</b>	0.52
Average K1 Ch3: 0.36								Average K2 Ch3: 0.29				
Average K1 Ch4: 0.41								Average K2 Ch4: 0.34				

- **κ1** refers to the kappa score for **SDS**, **κ2** refers to kappa score for all other **entities** together
- **Most** cases are **Fair** (0.21–0.40) to **moderate** (0.41–0.60) agreement
- **Few** cases are **slight** or **substantial** agreement
- (Ann3,Ann4) agree much more than any other annotator pair
- Agreement on **SDSs** is **higher** than agreement on **entities**.

# Results

Percentage agreement results for each entity type by token.

	Ann2				Ann3				Ann4			
	character	location	O	time	character	location	O	time	character	location	O	time
Ann1	0.24	0.15	0.99	0.8	0.28	0.15	0.99	0.26	0.33	0.05	0.99	0
Ann2					0.36	0.09	0.99	0.33	0.48	0.05	0.98	1
Ann3									<b>0.73</b>	<b>0.88</b>	0.98	1

- **O** refers to the **Outside** tag
- agreement for **Outside** tags will always be **high**, as most tokens are outside of any entity



## Summary of Results

---

- **Most** cases show **Fair** (0.21–0.40) to **moderate** (0.41–0.60) agreement
- **Few** cases show **slight** or **substantial** agreement
- Agreement on **SDSs** is **higher** than agreement on **entities**
- **Agreement scores** are **not** as **high** as we want
- Entity scores (k2) are dependent on agreement in SDS annotation (k1).
  - E.g. if one annotator annotates two SDSs where another annotates just one, the first annotator will have twice the number of time, location and character annotations, since these entity types are to be annotated for each SDS.





# Discussion

**Two** underlying causes of the observed disagreement:

1. Lack of **understanding** of the guidelines, the task and of English usage
2. Lack of **clarity** or specificity in the guidelines.

These are often not easy to distinguish

# Discussion

Examples of lack of understanding of the guidelines:

- **Two** distinct **location** entities tagged in the same **SDS**

Examples of lack of understanding English idioms and usage:

- **Idioms:** *How on earth had he* → (location)
- **general English:** *Sorry, old chap, had an attack of the wobbles.*  
*Dashed embarrassing* → (character)

# Discussion

Examples of lack of clarity and detail in the guidelines :

- Definite **articles**: *the stone ages* vs. *stone ages*
- *Do you not have good fabrics in the future?* → (time)
- Confusion regarding tagging **characters**:
  - annotate the fullest form?
  - the first mention?
  - or every mention?

# Discussion

Examples of lack of clarity and detail in the guidelines :

- Scene **transition** segments: e.g. *and soon I emerged back into the corridor looking like a new man*
  - Should this be annotated as belonging to the preceding or succeeding scene?
  - or is it a scene in its own right? or part of no scene?
  - or a new “scene transition” element?
- Scene change with minor character changes
  - If a very minor character leaves does this count as a change of scene?

# Conclusion and Future Work

- We have presented a **pilot annotation experiment** in which annotators were asked to use SceneML
  - The task is feasible, but we have learned several **changes** need to be made to:
    - the **annotation process**
    - the task **guidelines**
- in order to improve inter-annotator agreement

# Conclusion and Future Work

- Changes:
  - annotators should be better **trained**
  - annotators should do a **trial** annotation exercise, then be **filtered** from the annotator pool if their English language understanding or understanding of the task is too **poor**
  - the **guidelines** should be refined to reduce confusion

# Conclusion and Future Work

- Once agreement between annotators has been assured at a higher level we will :
  - extend the annotation to include **all SceneML** elements
  - apply the scheme to a much wider range of texts/media:
    - historical and contemporary fiction for adults, as well as biography
    - non-textual narrative genres, such as film
  - investigate **automating** the process of **scene annotation** after getting sufficient **manually annotated** data
  - investigate applying the scheme to languages other than English

# Thank you Questions?



✉ [ttalrashid1@sheffield.ac.uk](mailto:ttalrashid1@sheffield.ac.uk)

🐦 @AlrashidTarfah