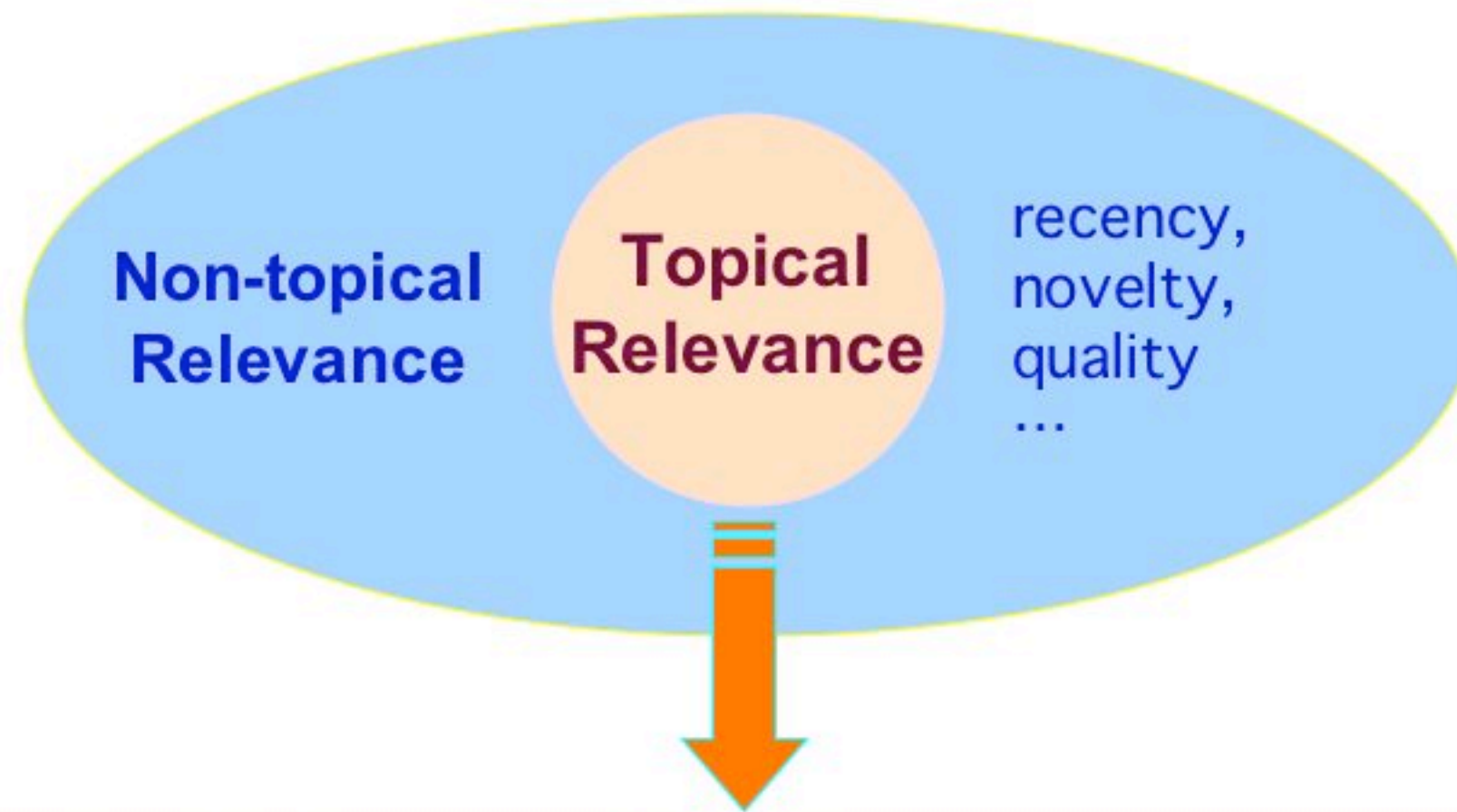


Evidence-based Interpretation of Topical Relevance Types: Towards A Richer Understanding of Topical Relevance

Xiaoli Huang and Dagobert Soergel

This work has been supported in part by NSF IIS Award 0122466.

Relevance



Study Goals

- Understand different ways in which evidence supports an argument
- Use the evidence framework to define and illustrate different types of topical relevance
- Apply to the design of IR systems and to information searching and use

Related Theories

- Philosophy and logic: Theory of evidence and inferential reasoning
- Communication: Relevance theory
- History: Evidence in historical reasoning
- Law: Law of evidence
- Medicine: Evidence-based medicine

Study Context – MALACH

Survivors of the Shoah Visual History Foundation
52,000 Holocaust survivor and liberator testimonies

MALACH Test Collection

- 75 total topics
 - 400 testimonies
 - 20,000 topical segments defined by cataloger
- Example Topics:
- Food in Auschwitz
 - Music in the Holocaust
 - Ghetto Life

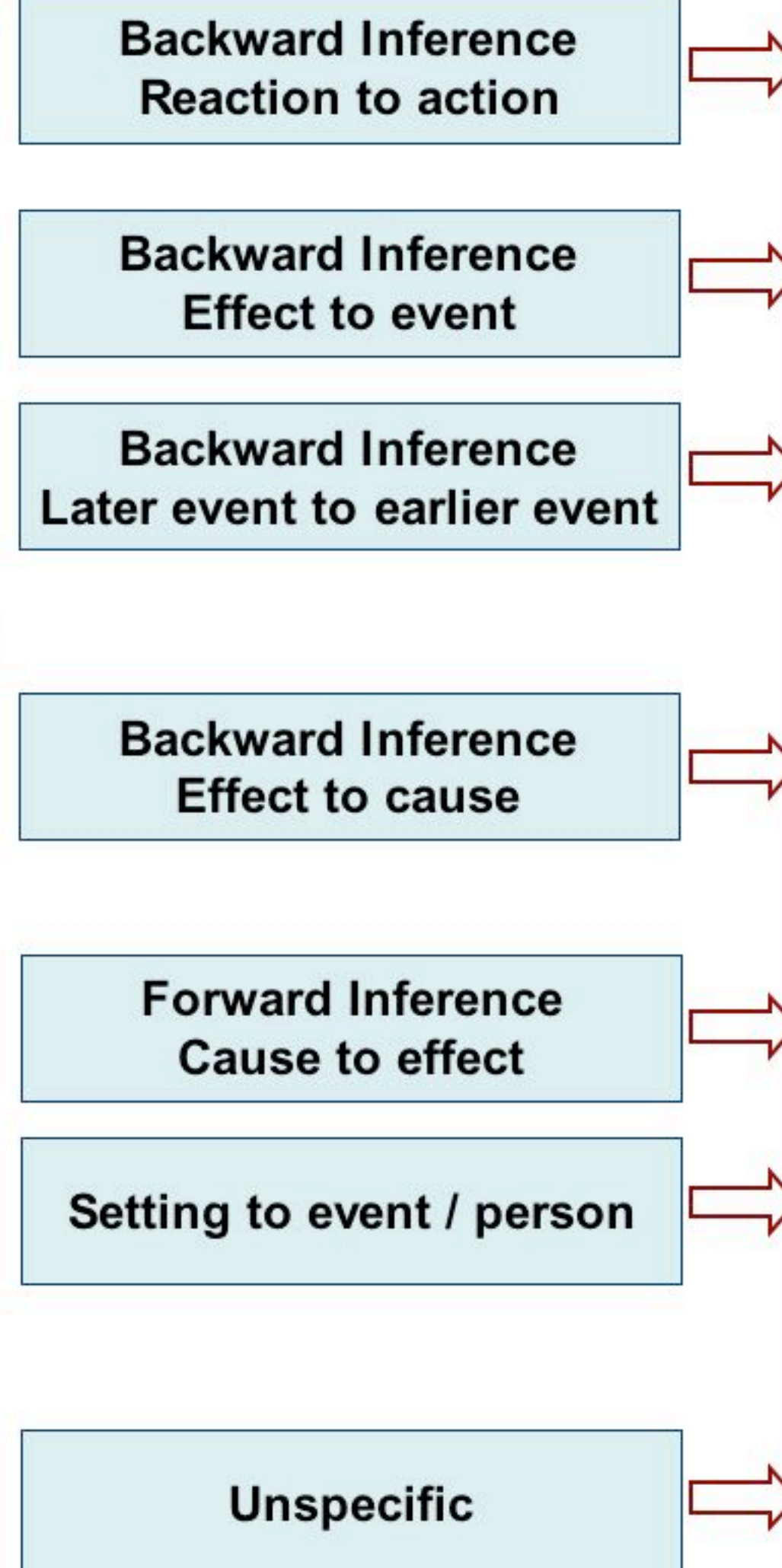
Relevance Assessments 2003 - 2005

- Nine history graduate student assessors
- 27,000 topic-segment pairs assessed with justifications

Topical Relevance Definitions

- Direct:** Direct evidence for what the user asks for
Talks about food given to Auschwitz inmates
- Indirect:** Provides indirect evidence on the topic; data from which one could infer something about the topic
Talks about seeing emaciated people in Auschwitz
- Context:** Provides background/context for topic; sets the stage for what user asks for; sheds additional light on topic
Talks about physical labor of Auschwitz inmates
- Comparison:** Provides information on similar/parallel situation or on contrasting situation
Talks about food in the Warsaw ghetto

Indirect



Topic Evidence

- Sexual experiences during the Holocaust** Survivor discusses how the women feared being raped. This indirectly indicates that much rape/sexual molestation was happening.
- Nazi eugenics policy** Survivor talks about how even now medical treatment still induced fear of medical experiments.
- Nazi theft in Eastern Europe** Survivor describes forced labor of sorting seized property (maybe just clothes or Jewish ritual objects). She does not specifically say that property was seized by the Nazis. From what happened later, sorting property, we infer what happened earlier, seizure of property by the Nazis.
- Bulgaria saved its Jews?** Survivor talks about friendship with surrounding non-Jews prewar and indicates that they lived in harmony in Bulgaria until the alliance with Germany. We infer that general populace and the government in Bulgaria was not conducive to anti-Semitism.
- Counterfeiting in Sachsenhausen** Survivor of the Sachsenhausen camp describes his school accounting classes. Because of this expertise, he may have been selected for counterfeiting.
- Eichmann witnesses** Segment provides location information for Eichmann without mentioning the location is related to Eichmann's activity. From typical locations where Eichmann actively carried out his rescue activities, we infer events happened in these locations at the time might have certain relations with Eichmann.
- Refugees assisted by Josef Schleich** Segment provides information on an escape across the Yugoslavian / Austrian border and may detail "a smuggler" in Styria, but does not name Schleich. During 1938-1942, Josef Schleich helped Jewish refugees escape from Nazi Germany and occupied countries via the Austrian-Yugoslavian border. Based on the time and location provided, the "smuggler" mentioned might well be Josef Schleich, though we do not have 100% certainty to make the claim.

(Huang & Soergel, ASIST 2004)

Topical Relationship Type	On topic	Relationship between overall document topic & overall user need topic	Diagram Illustrating Relationship ¹	Relatedness to the topic	Specificity of information	Richness of relevant information
Direct	Directly	Matching		High	High	High ~ medium
Indirect	After inference	Matching		High ~ medium	High	High ~ medium
Context	No	Surrounding		Low ~ medium	Low	Low ~ medium
Comparison	Partially	Comparing/contrasting		Low ~ medium	High ~ medium	Medium ~ low

¹ In the diagram, the document topic is black; the user need topic is white