



INTRODUCTION

- Radiology reports contain rich descriptions of clinically important findings and medical devices that are oftentimes referred to multiple times
- Tracking (or identifying coreferences) of radiological finding across reports is less studied unlike information extraction¹
- Introduce an annotated cross-document coreference resolution (CDCR) dataset of 60 MIMIC-III² patients to trac the same findings and devices across reports (5872) mentions with 2292 mention chains)
- Employ rule-based and transformer language-based systems to identify cross-report coreferences

ANNOTATION PROCESS

- Assign the same mention identifier to all mentions across reports representing the course of a finding or device
- Course of a finding (1) initial detection, (2) improved, worsened, & (3) no longer detected
- Course of a device (1) insertion, (2) device position status – normal/abnormal, and (3) removal
- Example (both aneurysms refer to the same entity):
 - Report 1: Questionable aneurysm at right posterior communicating artery
 - Report 4: Small aneurysm of size 2.5 mm arises at the origin of posterior communicating artery

								RE	SUL	TS					
Methods			MUC		B ³		CEAF _e		CONLL B		BLANC				
		Р	R	F1	Р	R	F1	Р	R	F1	F1	Р	R	F1	Tab 5. CDCR
String match		80.18	70.36	74.95	83.52	70.75	76.6	61.88	76.46	68.4	73.32	78.52	71.6	74.44	performances
BERT (mention	ns)	68.56	95.31	79.65	40.84	86.57	55.23	76.53	23.46	35.84	56.91	53.39	77.18	50.29	•
BERT (mention	ns + contex	kt) 67.46	92.58	77.87	32.72	85.7	46.48	76.12	18.97	29.99	51.45	49.79	67.92	39.13	
					Mentio	on pairs	Cor	respondir	ng sentenc	es				Category	Reason
Model Mentions + Context	P (%) 44.83 52.76	R (%) 85.89 86.3	F1 58.91 65.49	Acc 95.53 96.61	NG tube; NG tubeReport-5 Compared with prior radiograph, an NG tube has been withdrawn and there is significant dilatation of the colon lying just below the right hemidiaphragm; Report-10 An NGFPMore deep understan quired (e.g., "withdre cotes that the NG tube				More deep understanding of context is re- quired (e.g., " <i>withdrawn</i> " in Report-5 indi- cates that the NG tube in Report-10 is dif-						
Tab 6. BERT classifier results				lts	Report-3 The grayscale ultrasound of the veins of the upperthrombus; throm- busextremities demonstrated filling defect in the right cephalic vein at the level of the antecubital fossa consistent with <i>thrombus</i> ; Report-13 No intraluminal <i>thrombus</i> is identified						ferent from the first one). Sufficient contex- tual information is not incorporated into the models				
					collaps tases	e; atele	Rep c- with Ther well	ort-1 The left lowe re is cardio as atelect	ere is incr r lobe <i>coll</i> omegaly w asis in the	eased retr lapse and/ vith atelect left lower	rocardiac dens for consolidati tases in the left lobe.	sity, cons on; Repo ft upper lo	istent o rt-20 obe as	FN	More domain knowledge understanding is required to link the
Tab 7. BERT error analysis hen hen				hemorr hemato	hage; ma	Report-1 There is no intraparenchymal <i>hemorrhage</i> identified; Report-6 There is a small left frontal subdural <i>hematoma</i> , slightly larger than prior CT studies						conclated mulligs			

A Cross-document Coreference Dataset for Longitudinal Tracking across Radiology Reports

Surabhi Datta, Hio Cheng Lam, Atieh Pajouhi, Sunitha Mogalla, Kirk Roberts

The School of Biomedical Informatics |The University of Texas Health Science Center at Houston {surabhi.datta,kirk.roberts}@uth.tmc.edu

	_					Evenenia Battante		
	E	xample: Pat i #1 - 2197-04-2	i ent 1 5 04:08:00)	Re	Example: Patient 2 port #1 - 2111-12-14 19:25:0	0	
		ry edema has	progresse	ed	FIN The	NDINGS: ere is an NG tube present wit	th the	
	obscuring	g more focal a	areas of		tip	in the body of the stomach.		
	puimonai					• • •		
	Report #	2 - 2197-04-2	6 13:33:00		Re	port #2 - 2111-12-15 14:31:0	0	
	CHEST:				En	dotracheal tube terminates		
	chest x-r	edema prese	nt on the p	prior	ap	proximately 5.5 cm above the	level	
					En	iteric tube terminates within t	he	
		e in mediastina	al widening	g	sto	omach.		
	and perif	nilar <mark>edema</mark> .				• • •		
		•••			Re	port #3 - 2111-12-17 19:06:0	0	
	Report #	#3 - 2197-04-2	7 06:01:00	2	AP	UPRIGHT CHEST:		
	FINDING	is:	ie elisabele		ET	T has been removed.		
	improved	niiar naziness 1 since prior si	s is slightly tudy.	, 		• • •		
		•••			Т	racking of the same devic	e 1,	
	Trackir	ng of the san	ne findin	g		device 2		
	Fig 1.	Irackin	g eder	ma, NG	tube	& Endotracheal t	ube	
				ANN(DTA.	TION CHALI	LENG	
	CT HEAD	W/O CONTR	AST				Difficulty	,
	Findings:							
					— •			
	 The right	frontal fractur	e is associ	iated with a	Tab	3. Context	Simple (s	SV
	 The right focal lentife roughly 8 1	frontal fractur orm extra-axia mm in thickne	e is associ l hematom ess and 3.5	iated with a a measuring cm in maxi-	Tab depe	3. Context endency	Simple (sonymous)	sy
	 The right focal lentife roughly 8 mal transve This demo	frontal fractur orm extra-axia mm in thickne erse dimension	e is associ l hematom ess and 3.5 n.	iated with a a measuring cm in maxi-	Tab depe	3. Context endency	Simple (sonymous)	sy
	 The right focal lentife roughly 8 mal transve This demo portion, ar	frontal fractur orm extra-axia mm in thickne erse dimension onstrates a rela nteriorly, which	e is associ l hematom ess and 3.5 n. atively low h may rep	iated with a a measuring cm in maxi- v-attenuation resent acute,	Tab depe	3. Context endency	Simple (sonymous)	sy
	 The right focal lentife roughly 8 mal transve This demo portion, ar non-clotted This collec	frontal fractur orm extra-axia mm in thickne erse dimension onstrates a rela nteriorly, which l blood. tion may be b	e is associ l hematom ess and 3.5 n. atively low h may rep ounded by	iated with a a measuring cm in maxi- v-attenuation resent acute, v the coronal	Tab depe	3. Context endency	Simple (sonymous)	sy
	 The right focal lentife roughly 8 mal transve This demo portion, ar non-clotted This collect suture, and	frontal fractur orm extra-axia mm in thickne erse dimension onstrates a rela- nteriorly, which blood. tion may be b therefore lie in	e is associated in the matom atively low h may repart ounded by	iated with a a measuring cm in maxi- v-attenuation resent acute, v the coronal ural space.	Tab depe	3. Context endency	Simple (sonymous) Simple (clinical knowledg	е)
	 The right focal lentife roughly 8 mal transve This demo portion, ar non-clotted This collect suture, and There is me cupying th	frontal fractur orm extra-axia mm in thickne erse dimension onstrates a rela teriorly, which blood. tion may be b therefore lie in oderate subara	e is associ l hematom ess and 3.5 n. atively low h may rep ounded by n the epidu chnoid her y subjacent	iated with a a measuring cm in maxi- v-attenuation resent acute, v the coronal tral space. morrhage oc- t sulci, which	Tab depe	3. Context endency	Simple (sonymous) Simple (clinical knowledg	e)
	 The right focal lentife roughly 8 mal transve This demo portion, ar non-clotted This collect suture, and There is me cupying th are slightly	frontal fractur orm extra-axia mm in thickne erse dimension onstrates a rela- nteriorly, which blood. tion may be b therefore lie in oderate subara e immediately flattened, due	e is associated in the material hematom ess and 3.5 h. And the structure of the subjacent of the mass of the mass of the structure of the stru	iated with a a measuring cm in maxi- v-attenuation resent acute, the coronal aral space. morrhage oc- t sulci, which s effect of the	Tab depe	3. Context endency	Simple (sonymous) Simple (clinical knowledg Moderate	e)
	The right focal lentife roughly 8 mal transve This demo portion, ar non-clotted This collect suture, and There is mo cupying th are slightly hematoma.	frontal fractur orm extra-axia mm in thickne erse dimension onstrates a rela- nteriorly, which blood. tion may be b therefore lie in oderate subara e immediately flattened, due	e is associated in the epiduation of the epiduation of the epiduation of the epiduation of the mass of	iated with a a measuring cm in maxi- v-attenuation resent acute, v the coronal aral space. morrhage oc- t sulci, which s effect of the	Tab depe	3. Context endency	Simple (sonymous) Simple (clinical knowledg Moderate (clinical knowledg	e)
	 The right focal lentife roughly 8 mal transve This demo portion, ar non-clotted This collect suture, and There is me cupying th are slightly hematoma. No signific at the corre	frontal fractur orm extra-axia mm in thickne erse dimension onstrates a rela- nteriorly, which blood. tion may be b therefore lie in oderate subara e immediately flattened, due	re is associated in the material of the materi	iated with a a measuring cm in maxi- v-attenuation resent acute, v the coronal aral space. morrhage oc- t sulci, which s effect of the a is identified poral fracture	Tab depe	3. Context endency	Simple (sonymous) Simple (clinical knowledg Moderate (clinical knowledg	e)
	The right focal lentife roughly 8 mal transve This demo portion, ar non-clotted This collect suture, and There is me cupying th are slightly hematoma. No signific at the corre- site, though	frontal fractur orm extra-axia mm in thickne erse dimension onstrates a rela- nteriorly, which blood. tion may be b therefore lie in oderate subara e immediately flattened, due ant extra-axial esponding left in there is subar	re is associated thematom ass and 3.5 atively low h may repared ounded by n the epidu chnoid her to the mass hematoma frontotemp achnoid her	iated with a na measuring cm in maxi- v-attenuation resent acute, v the coronal ural space. morrhage oc- t sulci, which s effect of the a is identified poral fracture emorrhage in	Tab depe	3. Context endency	Simple (sonymous) Simple (clinical knowledg Moderate (clinical knowledg	e)
	 The right focal lentife roughly 8 mal transve This demo portion, ar non-clotted This collect suture, and There is me cupying th are slightly hematoma. No signific at the corres site, though the sulci in	frontal fractur orm extra-axia mm in thickne erse dimension onstrates a rela- nteriorly, which blood. tion may be b therefore lie in oderate subara e immediately flattened, due ant extra-axial esponding left in there is subara this region .	re is associal hematom ess and 3.5 atively low h may reparation ounded by n the epidu chnoid her subjacent to the mass hematoma frontotemp cachnoid he	iated with a a measuring cm in maxi- v-attenuation resent acute, v the coronal aral space. morrhage oc- t sulci, which s effect of the a is identified boral fracture emorrhage in	Tab depe	3. Context endency Tab 4. Domain	Simple (sonymous) Simple (clinical knowledg Moderate (clinical knowledg	e)
	 The right focal lentife roughly 8 mal transve This demo portion, ar non-clotted This collect suture, and There is me cupying th are slightly hematoma. No signific at the corre- site, though the sulci in Impression 	frontal fractur orm extra-axia mm in thickne erse dimension onstrates a rela- nteriorly, which blood. tion may be b therefore lie in oderate subara e immediately flattened, due ant extra-axial esponding left in there is subara this region .	re is associal hematom ess and 3.5 atively low h may rep ounded by n the epidu chnoid her subjacent to the mass hematoma frontotemp achnoid he	iated with a na measuring cm in maxi- v-attenuation resent acute, the coronal and space. morrhage oc- sulci, which s effect of the a is identified boral fracture emorrhage in	Tab depe kr	3. Context endency Tab 4. Domain nowledge reliance	Simple (sonymous) Simple (clinical knowledg Moderate (clinical knowledg Complex (clinical	e)
	The right focal lentife roughly 8 mal transve This demo portion, ar non-clotted This collect suture, and There is me cupying th are slightly hematoma. No signific at the corre- site, though the sulci in Impression	frontal fractur orm extra-axia mm in thickne erse dimension onstrates a rela- nteriorly, which blood. tion may be b therefore lie in oderate subara e immediately flattened, due ant extra-axial esponding left in there is subar this region.	re is associal hematom ess and 3.5 atively low h may repared ounded by n the epidu chnoid her subjacent to the mass hematoma frontotemp achnoid he	iated with a a measuring cm in maxi- v-attenuation resent acute, v the coronal and space. morrhage oc- t sulci, which s effect of the a is identified boral fracture emorrhage in	Tab depe kr	3. Context endency Tab 4. Domain nowledge reliance	Simple (sonymous) Simple (clinical knowledg Moderate (clinical knowledg Complex (clinical knowledg	e)
	The right focal lentife roughly 8 mal transve This demo portion, ar non-clotted This collect suture, and There is me cupying th are slightly hematoma. No signific at the corres site, though the sulci in Impression Associated scribed ab right front 	frontal fractur orm extra-axia mm in thickne erse dimension onstrates a rela- nteriorly, which blood. tion may be b therefore lie in oderate subara e immediately flattened, due ant extra-axial esponding left in there is subara this region.	re is associal hematom ess and 3.5 a. atively low h may rep ounded by n the epidu chnoid her subjacent to the mass hematoma frontotemp achnoid he hemorrhag ssible sma htotempora	iated with a na measuring cm in maxi- v-attenuation resent acute, v the coronal ural space. morrhage oc- t sulci, which s effect of the a is identified boral fracture emorrhage in ge at sites de- ll associated al contusions	Tab depe kr	3. Context endency Tab 4. Domain nowledge reliance	Simple (sonymous) Simple (clinical knowledg Moderate (clinical knowledg Complex (clinical knowledg	e)
	The right focal lentife roughly 8 mal transve This demo portion, ar non-clotted This collect suture, and There is me cupying th are slightly hematoma. No signific at the corres site, though the sulci in Impression Associated scribed ab right front	frontal fractur orm extra-axia mm in thickne erse dimension onstrates a rela- neteriorly, which blood. tion may be b therefore lie in oderate subara e immediately flattened, due ant extra-axial esponding left in there is subar this region. h : subarachnoid ove , with por al and left fron	re is associal hematom ess and 3.5 a. atively low h may repared ounded by n the epidu chnoid her subjacent to the mass hematoma frontotemp achnoid he sachnoid her	iated with a a measuring cm in maxi- v-attenuation resent acute, v the coronal and space. morrhage oc- sulci , which s effect of the a is identified boral fracture emorrhage in ge at sites de- ll associated al contusions	Tab depe	3. Context endency Tab 4. Domain nowledge reliance	Simple (sonymous) Simple (clinical knowledg Moderate (clinical knowledg Complex (clinical knowledg	e)
	The right focal lentife roughly 8 mal transve This demo portion, ar non-clotted This collect suture, and There is me cupying th are slightly hematoma. No signific at the corres site, though the sulci in Impression Associated scribed ab right front	frontal fractur orm extra-axia mm in thickne erse dimension onstrates a rela- neteriorly, which blood. tion may be b therefore lie in oderate subara e immediately flattened, due ant extra-axial esponding left in there is subar this region. n: subarachnoid ove , with pot al and left fron	re is associal hematom ess and 3.5 a. atively low h may repared ounded by n the epidu chnoid her subjacent to the mass hematoma frontotemp achnoid he sachnoid he	iated with a na measuring cm in maxi- v-attenuation resent acute, v the coronal ral space. morrhage oc- t sulci , which s effect of the a is identified boral fracture emorrhage in ge at sites de- ll associated al contusions	Tab depe	3. Context endency Tab 4. Domain nowledge reliance	Simple (sonymous) Simple (clinical knowledg Moderate (clinical knowledg Complex (clinical knowledg	e)
Fe	The right focal lentife roughly 8 mal transve This demo portion, ar non-clotted This collect suture, and There is me cupying th are slightly hematoma. No signific at the corres site, though the sulci in Impression Associated scribed ab right front TSS	frontal fractur orm extra-axia mm in thickness erse dimension onstrates a rela- nteriorly, which blood. tion may be b therefore lie in oderate subara e immediately flattened, due ant extra-axial esponding left in there is subara this region. n: subarachnoid ove, with pos al and left from al and left from CONLL F1	re is associal hematom ess and 3.5 atively low h may rep ounded by n the epidu chnoid her subjacent to the mass hematoma frontotemp achnoid he sachnoid her achnoid her hematoma frontotemp achnoid her sachnoid her hemorrhag	iated with a a measuring cm in maxi- v-attenuation resent acute, v the coronal aral space. morrhage oc- t sulci, which s effect of the a is identified boral fracture emorrhage in ge at sites de- ll associated al contusions BLANC R	Tab depe	3. Context endency Tab 4. Domain nowledge reliance	Simple (sonymous) Simple (clinical knowledg Moderate (clinical knowledg Complex (clinical knowledg	e)
F _e	The right focal lentife roughly 8 mal transve This demo portion, ar non-clotted This collect suture, and There is me cupying th are slightly hematoma. No signific at the corres site, though the sulci in Impression Associated scribed ab right front TSS	frontal fractur orm extra-axia mm in thickne erse dimension onstrates a rela- nteriorly, which blood. tion may be b therefore lie in oderate subara e immediately flattened, due ant extra-axial esponding left in there is subar a this region. n: subarachnoid ove, with por al and left from al and left from 73.32	e is associal hematom ess and 3.5 atively low h may rep ounded by n the epidu chnoid her subjacent to the mass hematoma frontotemp achnoid he hemorrhag ssible sma htotempora	iated with a na measuring cm in maxi- w-attenuation resent acute, y the coronal and space. morrhage oc- t sulci, which s effect of the a is identified boral fracture emorrhage in ge at sites de- ll associated al contusions BLANC R 71.6	Tab depe kr	3. Context endency Tab 4. Domain nowledge reliance	Simple (sonymous) Simple (clinical knowledg Moderate (clinical knowledg Complex (clinical knowledg	e)
F _e	The right focal lentife roughly 8 mal transve This demo portion, ar non-clotted This collect suture, and There is me cupying th are slightly hematoma. No signific at the corres site, though the sulci in Associated scribed ab right front TS 68.4 35.84	frontal fractur orm extra-axia mm in thickne erse dimension onstrates a rela- nteriorly, which blood. tion may be b therefore lie in oderate subara e immediately flattened, due ant extra-axial esponding left in there is subara this region. n: subarachnoid ove, with pos al and left fron al and left fron ove, with pos al and left fron 56.91	re is associal hematom ess and 3.5 atively low h may rep ounded by n the epidu chnoid her subjacent to the mass hematoma frontotemp achnoid he sible sma hemorrhag ssible sma hemorrhag ssible sma hemorrhag ssible sma	iated with a na measuring cm in maxi- w-attenuation resent acute, whe coronal and space. morrhage oc- t sulci, which s effect of the a is identified boral fracture emorrhage in ge at sites de- ll associated al contusions BLANC R 71.6 77.18	Tab depe 50.29	3. Context endency Tab 4. Domain nowledge reliance Tab 5. CDCR performances	Simple (sonymous) Simple (clinical knowledg Moderate (clinical knowledg Complex (clinical knowledg	e)
F _e	The right focal lentife roughly 8 mal transve This demo portion, ar non-clotted This collect suture, and There is me cupying th are slightly hematoma. No signific at the corres site, though the sulci in Associated scribed ab right front TSS F1 68.4 35.84 29.99	frontal fractur orm extra-axia mm in thickne erse dimension onstrates a rela- nteriorly, which blood. tion may be b therefore lie in oderate subara e immediately flattened, due ant extra-axial esponding left in there is subar this region. n: subarachnoid ove, with pot al and left fron al and left fron 51.45	re is associal hematom ess and 3.5 atively low h may rep ounded by n the epidu chnoid her subjacent to the mass hematoma frontotemp achnoid he hemorrhag ssible sma ntotempora hemorrhag ssible sma ntotempora	iated with a measuring cm in maxi- cm in maxi- v-attenuation resent acute, y the coronal ral space. morrhage oc- t sulci, which s effect of the s effect of the sociated poral fracture emorrhage in sociated al contusions ge at sites de- ll associated al contusions BLANC R 71.6 77.18 67.92	Tab depe 4 5 1 1 1 1 1 1 1 1 1 1 1 1 1	3. Context endency Tab 4. Domain nowledge reliance Tab 5. CDCR performances	Simple (sonymous) Simple (clinical knowledg Moderate (clinical knowledg Complex (clinical knowledg	e)

					DATA
	Item			Count	Fig 2. Cov
CS	Avg no. o	f reports j	per patient	10.6	in
Sti	Total repo	orts		638	
ati	Avg no. o	f tokens p	er report	244.7	
St	Min no. o	f mention	chain per patient	8	
et	Max no. c	of mentior	n chain per patient	110	Fig 2 Imaging may
as(Total men	tion chair	IS	2292	rig J. Intaging mod
ata	Total sing	leton men	tion chains	1102	in mention
Õ	Longest c	hain lengt	th	53	
-	Avg chain	length (e	excluding singletons)	4	Size
	Avg no. o	f tokens p	per mention	1.44	tion
La	Total entit	ties (radio	logical finding)	4978	ersec
•	Total entit	ties (medi	cal device)	894	
Finc	ling	Count	Device	Count	
effu	sion	398	tip	144	OTI
pneı	umothorax	238	ng tube	103	
fracture		229	endotracheal tube	101	
opac	city	180	chest tube	42	
atele	ectasis	176	swan-ganz catheter	36	
-	Tab 2 . T	op five	frequent menti	ons	1000 0

GE	S
ty	Example pairs
(syn- s)	Some degree of dissociation as we as lateral displacement of the ossic lar chain; Complex fracture of the la temporal bone with evidence of later displacement and disruption of the la ossicular chain
lge)	There is a left parietovertex soft tiss swelling; There is extensive left supr and periorbital soft tissue edema
te l lge)	There is new patchy opacity at the left lung base, which may represent resolving postoperative atelect sis with effusion, but <i>pneumonia</i> can not be excluded; New retrocardiac con lapse/consolidation and bilateral effects sions
x l lge)	There is mild prominence of the pumonary vascular markings without overt evidence for failure; In the is terval, there is increased interstition edema and small-moderate bilater pleural effusions.

Method 1 - Rule-based:

String matching-based

Method 2 - BERT-based:

- Given a mention pair, BERT classifier to predict if two mer coreferences
- BERT input sequence: 1) onl mention spans, 2) additional (anatomy and modifier)
- Output generated by BERT for mention pairs for a patient is to predict the final mention ch

DISCUSSION

- CDCR is relatively under-explored in clinical domain³
- Further model improvement incorporate rich radiology-specific domain knowledge into the models; allow models to learn broad crossreport context
- Explore cross-document language modeling⁴ (effective for multi-document downstream tasks)
- Use predicted mention spans to infer the mention chains instead of gold mentions





	•	We construct a new CDCR dataset to track radiological findings and devices across reports
as a binary ntions are	•	We highlight major challenges (from annotation & model development perspectives)
ly the information	•	System performances are low to moderate, thus the dataset can be
or all combined hains		leveraged by researchers to develop more advanced methods for radiology CDCR in the future

ACKNOWLEDGEMENTS

National Institute of Biomedical Imaging and Bioengineering (R21EB029575), Patient-Centered Outcomes Research Institute (ME-2018C1-10963), and Cancer Prevention and Research Institute of Texas Predoctoral Fellowship (CPRIT RP210045)

REFERENCES

Steinkamp, J. M., Chambers, C., Lalevic, D., Zafar, H. M., and Cook, T. S. (2019). Toward Complete Structured Information Extraction from Radiology Reports Using Machine Learning. Journal of Digital Imaging, 32(4):554–564. Johnson, A. E., Pollard, T. J., Shen, L., Lehman, L. W. H., Feng, M., Ghassemi, M., Moody, B., Szolovits, P., Anthony Celi, L., and Mark, R. G. (2016). MIMIC-III, a freely accessible critical care database. Scientific Data, 3:160035.

Wright-Bettner, K., Palmer, M., Savova, G., de Groen, P., and Miller, T. (2019). Cross-document coreference: An approach to capturing coreference without context. In Proceedings of the Tenth International Workshop on Health Text Mining and Information Analysis (LOUHI 2019), pages 1–10.

Caciularu, A., Cohan, A., Beltagy, I., Peters, M., Cattan, A., and Dagan, I. (2021). CDLM: Cross-Document Language Modeling. In Findings of the Association for Computational Linguistics: EMNLP 2021, pages 2648–2662.