

Fall 2022 Semester End Report

Todd Morrill
Columbia University
tm3229@columbia.edu

December 2022

1 Summary

This was my first semester working in Professor Kathleen McKeown’s lab and in particular on the DARPA Computational Cultural Understanding (CCU) [DAR21] project. Internally at Columbia, this project is known as Cross-Cultural Harmony through Affect and Response Mediation (CHARM). My primary areas of focus were: 1) exploring novel approaches for change point prediction (i.e. TA1.3) using circumplex theory and 2) ramping up on the integration role.

Change Point Data Annotation My primary contributions were a data annotation scheme and mini experiment to assess the predictive power of circumplex theory with respect to the change point prediction task.

Integration My primary contributions were to prepare myself to assume the integration lead role next semester. In order to do that, I created a master inventory of all data releases along with useful metadata (e.g. modality, annotation status, etc.), participated in weekly integration meetings, and learned about the integration process by reading key documents (e.g. evaluation documentation, etc.).

2 Change Point Data Annotation

Annotated data is a crucial resource needed for developing performant change point prediction systems. One of my first tasks at the start of the semester was to create a change point labeling plan and scope out the cost required to annotate a small corpus [Mor22b]. This plan included a sample conversation (see figure 1) provided by LDC along with definitions for a variety of annotation tasks including communication change, sentiment, entrainment, and circumplex theory communicative behavior indicators. This plan was developed before LDC provided any annotated data and wasn’t pursued further once we had sufficient LDC annotated data.

Participant	Original Text	Translation	Time	Sentiment	Entrainment (topic)	Communicative Behavior Indicators (Circumplex)	Comm. change
A	想找个人聊聊~来吧	want to find someone to chat with ~ come on	2011-11-06 09:41:50	1	0	Informal	1
B	哈哈, 美国人	haha, American	2011-11-06 09:42:25	1	1	Emotionally expressive	1
B	啥时候回来呀?	<u>when</u> are you coming back?	2011-11-06 09:42:28	0	0	Emotionally expressive	0
A	12月19	December 19th	2011-11-06 09:42:	0	1	Direct	0

Figure 1: Sample annotations provided in the first change point labeling plan.

One of Columbia’s proposed approaches to the change point prediction task was to use circumplex theory as a proxy for the change point prediction task. Circumplex theory comes from the field of social psychology and aims to characterize interpersonal relations using a variety of social tags (see figure 2). The motivation for developing a circumplex theoretic system as a proxy for change point is as follows. If a circumplex theoretic system is easier to develop (e.g. because of existing data resources, etc.) and is highly correlated with the change point task, then we might improve the performance of our change point prediction system without the need for a large volume of change point prediction annotations.



Figure 2: Circumplex

In order to validate this approach, I developed a second data labeling plan focused on social relation tags [Mor22a]. Social relation tags characterize a person’s behavior throughout an interaction and include the following tagset: {Assured-Dominant, Gregarious-Extraverted, Warm-Agreeable, Unassuming-Ingenuous, Unassured-Submissive, Aloof-Introverted, Cold, Arrogant-Calculating}. Yanda Chen and Yukun Huang volunteered to annotate several change point videos using the circumplex theory social relation tags. I then compared their predictions to the ground truth change point annotations provided by LDC and found that the social relation tagging scheme did correlate with change point prediction task. In particular, Yanda and Yukun’s social relation tags found 100% of the LDC change point annotations, though there were false positives (see figure 3). My sense is that social relation tags have the potential to be even more fine-grained than the change point task. These fine-grained predictions can then be fed to a downstream system that can filter through these social relation tags and decide what’s useful for the change point prediction task.

Gold Standard Reference Annotations

Time	Speaker ID	Speaker Descriptor	Tag	Explanation
0:05	1	Woman, yellow jacket	Gregarious-Extraverted	Hosting the video, warm-greeting
1:10	2	Man, red shirt	Unassuming-Ingenuous	Everytime the woman asked a question, he gave an answer but he didn't bring up topic on his own topic, he "takes things as they come"
4:00	3	Man, white shirt, short sleeves	Gregarious-Extraverted	Talks to a lot of different people
4:08	4	Man, black shirt	Gregarious-Extraverted	The man spoke to the camera from far away and wanted to catch attention
4:29	5	Man, blue shirt	Gregarious-Extraverted	The man walked straight to the camera and wanted to be interviewed
6:00	6	Man, driving motorcycle, black jacket	Gregarious-Extraverted	Offered to introduce the mountain
6:20	7	Man, driving motorcycle, brown jacket	Arrogant-Calculating	Likes to contradict to the woman

Figure 3: Social relation

Throughout the semester, I met with and corresponded with Colin Leach, Professor of Psychology & Africana Studies from Barnard College, who provided helpful guidance and resources. Professor Leach shared a dictionary of English words tagged to 128 categories such as positive emotion, negative emotion, affect, etc. (see figure xx). One approach to social relation tagging is weak supervision, whereby we tag each word in a person's utterance to a particular social relation tag. We can then create a signature of the person's social relation tag through summary statistics (e.g. sum words used for each social relation tag).

30 Affect										31 Posemo										32 Negemo									
0	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	0	entertain	high	proving	wealthy	0	distast	ignorant	poor	tragic	30	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
1	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	1	entertain	high	proving	wealthy	1	distast	ignorant	poor	tragic	31	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
2	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	2	entertain	high	proving	wealthy	2	distast	ignorant	poor	tragic	32	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
3	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	3	entertain	high	proving	wealthy	3	distast	ignorant	poor	tragic	33	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
4	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	4	entertain	high	proving	wealthy	4	distast	ignorant	poor	tragic	34	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
5	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	5	entertain	high	proving	wealthy	5	distast	ignorant	poor	tragic	35	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
6	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	6	entertain	high	proving	wealthy	6	distast	ignorant	poor	tragic	36	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
7	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	7	entertain	high	proving	wealthy	7	distast	ignorant	poor	tragic	37	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
8	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	8	entertain	high	proving	wealthy	8	distast	ignorant	poor	tragic	38	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
9	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	9	entertain	high	proving	wealthy	9	distast	ignorant	poor	tragic	39	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
10	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	10	entertain	high	proving	wealthy	10	distast	ignorant	poor	tragic	40	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
11	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	11	entertain	high	proving	wealthy	11	distast	ignorant	poor	tragic	41	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
12	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	12	entertain	high	proving	wealthy	12	distast	ignorant	poor	tragic	42	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
13	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	13	entertain	high	proving	wealthy	13	distast	ignorant	poor	tragic	43	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
14	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	14	entertain	high	proving	wealthy	14	distast	ignorant	poor	tragic	44	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
15	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	15	entertain	high	proving	wealthy	15	distast	ignorant	poor	tragic	45	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
16	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	16	entertain	high	proving	wealthy	16	distast	ignorant	poor	tragic	46	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
17	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	17	entertain	high	proving	wealthy	17	distast	ignorant	poor	tragic	47	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
18	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	18	entertain	high	proving	wealthy	18	distast	ignorant	poor	tragic	48	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
19	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	19	entertain	high	proving	wealthy	19	distast	ignorant	poor	tragic	49	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
20	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	20	entertain	high	proving	wealthy	20	distast	ignorant	poor	tragic	50	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
21	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	21	entertain	high	proving	wealthy	21	distast	ignorant	poor	tragic	51	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
22	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	22	entertain	high	proving	wealthy	22	distast	ignorant	poor	tragic	52	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
23	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	23	entertain	high	proving	wealthy	23	distast	ignorant	poor	tragic	53	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
24	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	24	entertain	high	proving	wealthy	24	distast	ignorant	poor	tragic	54	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
25	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	25	entertain	high	proving	wealthy	25	distast	ignorant	poor	tragic	55	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
26	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	26	entertain	high	proving	wealthy	26	distast	ignorant	poor	tragic	56	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
27	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	27	entertain	high	proving	wealthy	27	distast	ignorant	poor	tragic	57	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
28	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	28	entertain	high	proving	wealthy	28	distast	ignorant	poor	tragic	58	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
29	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	29	entertain	high	proving	wealthy	29	distast	ignorant	poor	tragic	59	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
30	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	30	entertain	high	proving	wealthy	30	distast	ignorant	poor	tragic	60	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
31	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	31	entertain	high	proving	wealthy	31	distast	ignorant	poor	tragic	61	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
32	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	32	entertain	high	proving	wealthy	32	distast	ignorant	poor	tragic	62	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
33	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	33	entertain	high	proving	wealthy	33	distast	ignorant	poor	tragic	63	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
34	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	34	entertain	high	proving	wealthy	34	distast	ignorant	poor	tragic	64	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
35	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	35	entertain	high	proving	wealthy	35	distast	ignorant	poor	tragic	65	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
36	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	36	entertain	high	proving	wealthy	36	distast	ignorant	poor	tragic	66	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
37	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	37	entertain	high	proving	wealthy	37	distast	ignorant	poor	tragic	67	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
38	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	38	entertain	high	proving	wealthy	38	distast	ignorant	poor	tragic	68	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
39	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	39	entertain	high	proving	wealthy	39	distast	ignorant	poor	tragic	69	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
40	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	40	entertain	high	proving	wealthy	40	distast	ignorant	poor	tragic	70	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
41	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	41	entertain	high	proving	wealthy	41	distast	ignorant	poor	tragic	71	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
42	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	42	entertain	high	proving	wealthy	42	distast	ignorant	poor	tragic	72	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
43	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	43	entertain	high	proving	wealthy	43	distast	ignorant	poor	tragic	73	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
44	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	44	entertain	high	proving	wealthy	44	distast	ignorant	poor	tragic	74	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
45	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	45	entertain	high	proving	wealthy	45	distast	ignorant	poor	tragic	75	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
46	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	46	entertain	high	proving	wealthy	46	distast	ignorant	poor	tragic	76	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
47	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	47	entertain	high	proving	wealthy	47	distast	ignorant	poor	tragic	77	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
48	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	48	entertain	high	proving	wealthy	48	distast	ignorant	poor	tragic	78	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
49	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	49	entertain	high	proving	wealthy	49	distast	ignorant	poor	tragic	79	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
50	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	50	entertain	high	proving	wealthy	50	distast	ignorant	poor	tragic	80	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
51	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	51	entertain	high	proving	wealthy	51	distast	ignorant	poor	tragic	81	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
52	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	52	entertain	high	proving	wealthy	52	distast	ignorant	poor	tragic	82	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
53	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	53	entertain	high	proving	wealthy	53	distast	ignorant	poor	tragic	83	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
54	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	54	entertain	high	proving	wealthy	54	distast	ignorant	poor	tragic	84	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
55	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	55	entertain	high	proving	wealthy	55	distast	ignorant	poor	tragic	85	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
56	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	56	entertain	high	proving	wealthy	56	distast	ignorant	poor	tragic	86	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
57	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	57	entertain	high	proving	wealthy	57	distast	ignorant	poor	tragic	87	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
58	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	58	entertain	high	proving	wealthy	58	distast	ignorant	poor	tragic	88	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship
59	challenge	blat	turn	hopeless	lowest	pleasing	stir	trist	worship	59	entertain	high	proving	wealthy	59	distast	ignorant	poor	tragic	89	challenge	blat</							

commanders, etc.) how their cross-cultural interactions are going and provide guidance to them so they can achieve their objectives. In addition to combining all the systems, the integration lead interfaces with LDC to inventory data releases and NIST to comply with the evaluation procedures of the program.

This semester I created a master inventory of all data releases along with useful metadata such as the modality, annotation status, etc. (see figure 5). While creating this metadata inventory, I translated hundreds of LDC documents to English so that 1) English speakers could inspect the documents and 2) we could apply models trained on English data only. During this exercise, I accounted for all files that were transcribed and then worked with Sukrit Rao and Michael Picheny at New York University (NYU) to prioritize transcriptions for mini evaluation data. I shared this metadata artifact with the CHARM team internally at Columbia and it was immediately useful to those that used it (e.g. Tom Zollo). I also familiarized myself with Columbia’s proposal for the DARPA CCU project, NIST’s evaluation guidelines, and SRI’s resources (e.g. Confluence, Docker containers, message queues, etc.). I attended weekly integration meetings and recently presented an overview of the change point systems developed for the mini evaluation so that TA2 teams have more insight into how to use our systems.

release	file_uid	modality	url	emotion_count	valence_arousal_count	norms_count	changeoint_count	start	end	utterance_count	transcribed	translated
R2	M01003MO8	video	http	1.0	1.0	1.0	0.0	41.7	341.7			
R2	M01003MS5	video	http	1.0	1.0	1.0	0.0	114.8	414.8			
R2	M01003MSB	video	http	3.0	3.0	1.0	0.0	1016.0	1316.0			
R2	M01003MTK	video	http	3.0	3.0	1.0	1.0	0.0	300.0	93.0	TRUE	TRUE

Figure 5: Metadata

4 Next Steps

Circumplex theory: still need to come up with a mapping from the 128 categories to social relation tags

Heng Ji’s students’ annotations

GPT3 to generate a small labeled corpus

Integration: continue building out the data loader, host the data on Google Drive

References

- [DAR21] DARPA. *Creating AI-Enabled Cultural Interpreters to Aid Defense Operations*. May 2021. URL: <https://www.darpa.mil/news-events/2021-05-03a>.
- [Mor22a] Todd Morrill. *Circumplex Theory Data Annotation Instructions*. Nov. 2022. URL: <https://docs.google.com/document/d/1BHnwoCKFKPIXdu0ZZdoiNBSBsW-wYg3ubQjNe8Eakw/edit?usp=sharing>.
- [Mor22b] Todd Morrill. *Task 1.3 Data Labeling Plan*. Oct. 2022. URL: https://docs.google.com/document/d/1k_6FrhVcSk4CkeIqATBv5f4e-AiThJEgKDVg8RZydeI/edit?usp=sharing.