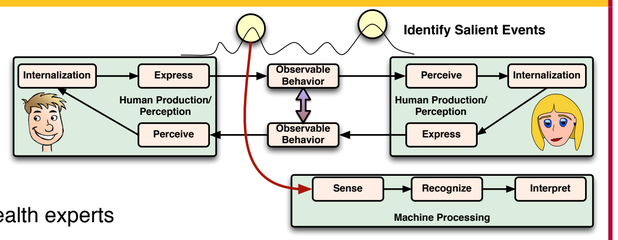
Quantitative Observational Practice in Family Studies (3 of 3)

Panayiotis (Panos) Georgiou, Shri Narayanan, Gayla Margolin, Brian Baucom, Matt Black, James Gibson, Nassos Katsamanis, Jeremy Lee, & Bo Xiao
Signal Analysis and Interpretation Laboratory (SAIL), Viterbi School of Engineering, University of Southern California, Los Angeles, CA 90089
georgiou@sipi.usc.edu

Abstract

Goal:

- Transform observational behavior analysis
- Through computational framework
- Modeling of emotionally-rich human interactions
- Signal processing and machine learning
- Existing family therapy data
- Alleviate the tedium of manual annotation
- Offer new analysis capabilities and empower the mental health experts



Significance: USA-10mil people receive psychotherapy every year and state of the art hasn't changed for decades

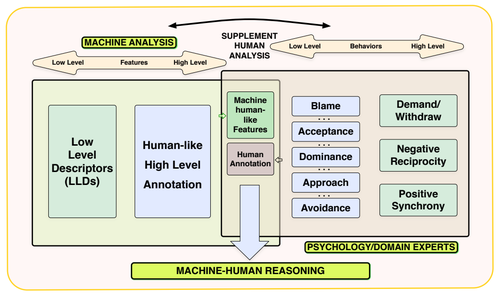
Approaches

+ This poster: [- Other two posters]

- Model interlocutors independently
- Model dynamics of interlocutors

+ Incorporate Saliency:

×Lexical, acoustic and visual modalities

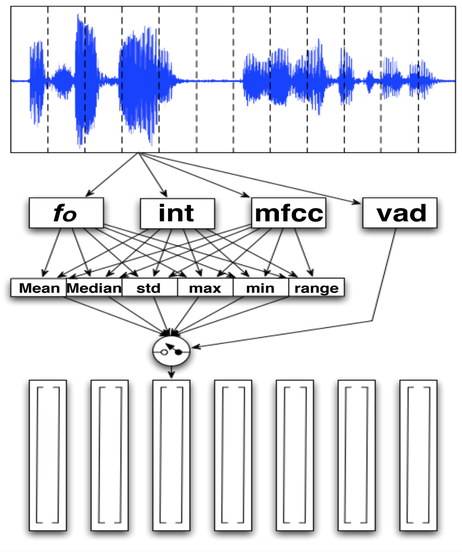


Saliency

- Couples' problem solving discussion are rated on a session level
- It is of interest to identify shorter-term events that influence evaluators' perceptions of the interaction
- These "salient" instances may help to inform both behavioral scientists
- We use multiple instance learning (MIL) to focus on local events in the couples' therapy sessions

What are the important bits?

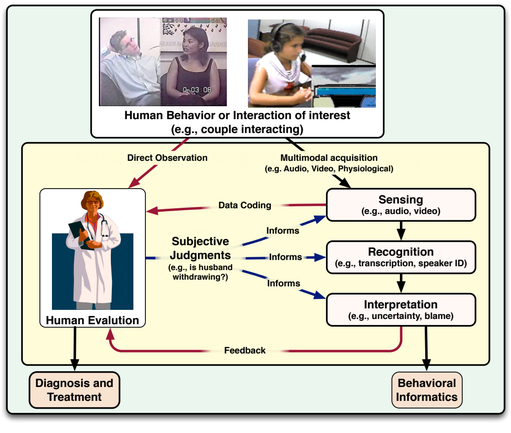
Audio Feature Extraction



Data

Couple Therapy Corpus

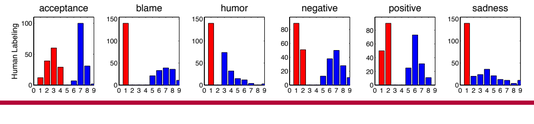
- 117 real distressed couples
- 10-minute dyadic interactions
- 596 sessions (96 hours)



Data used

Audio/Lexical and Visual subsets used

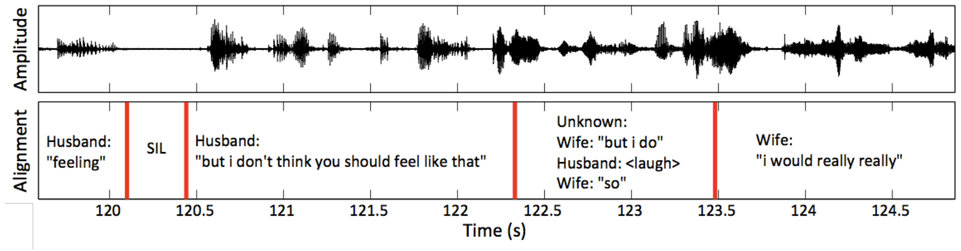
- Use top/bottom 20% for audio, lexical and 25% for video
- Choose subsets with acceptable audio/video qualities
- Used 6 codes with highest human agreement
- Some distributions skewed and not very separable



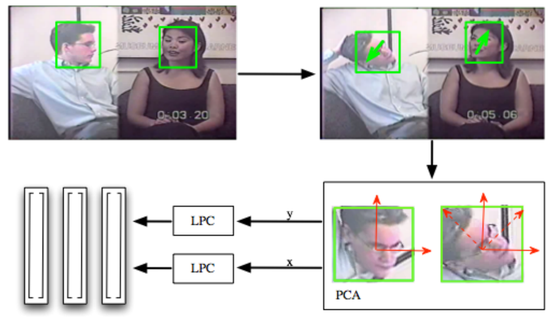
Multiple Instance Learning: Instances

Identify ***saliency*** through Multiple Instance Learning

- We consider each session a "bag" of "instances"
- Instances are varying-length speaker turns or equal-length windows
- Each instance conveys particular behaviors of interest with varying degrees
- MIL is a method for identifying the "salient instances", i.e., the local events that most greatly affect the final rating assigned to the session



Visual Feature Extraction

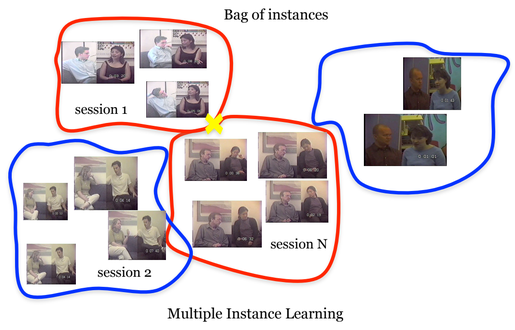


Lexical Feature Extraction

Behavior	Informative words
acceptance	UM, TOLD, NOTHING, MM, YES, EVERYTHING, ASK, MORE, (LAUGH), CAN'T
blame	NOTHING, EVERYTHING, YOUR, NO, SAID, ALWAYS, CAN'T, NEVER, MM, TOLD
humor	(LAUGH), TOPIC, GOOD, MISSING, COOL, TREAT, SEEMED, TRULY, ACCEPT, CASE
negative	TOLD, KIND, MM, MAYBE, NOTHING, UM, YOUR, NEVER, CAN'T, (LAUGH)
positive	UM, KIND, NOTHING, MM, GOOD, (LAUGH), TOLD, CAN'T, MEAN, WHY
sadness	ACTUALLY, ONCE, WEEK, GO, OKAY, STAND, CONSTANTLY, UP, ALREADY, WENT

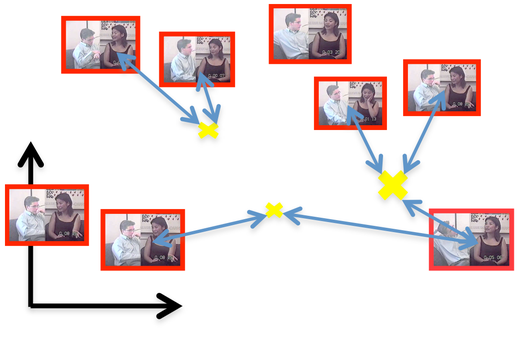
Session Level Feature Extraction

Salient feature identification:



Multiple Instance Learning

Distance of session features to salient prototype features



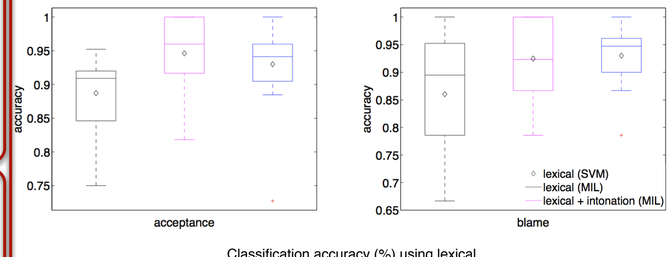
Behavioral classification through MIL

Configuration	Wife		Husband	
	DD-SVM	SVM	DD-SVM	SVM
acceptance	**73.6**	72.1	69.3	69.3
blame	77.1	**79.3**	**72.3**	71.4
positive	**74.3**	70.0	55.0	**58.6**
negative	75.0	**77.9**	**71.4**	70.0
sadness	**66.4**	57.1	**63.6**	62.9
humor	**52.9**	51.4	63.6	63.6

Classification accuracy (%) using audio with utterance level instances

behavior	audio	visual	fusion	
			early	late
acceptance	70.5	62.5	64.3	72.3
blame	69.4	57.4	70.4	71.3

Classification accuracy (%) using audio, visual, and audio-visual fusion with overlapping two second instances



Classification accuracy (%) using lexical, intonation, and lexical-intonation fusion features

Summary and Future work

- Explored saliency in MIL framework
- Explored saliency in multiple modalities
- Explored low-level instance features and deriving high-level session features
- Temporal dynamics of salient events for ***reactivity***
- Explore alternative measures for saliency, such as knowledge inspired signal cues (e.g., laughter, crying)

Citations, Acknowledgments

Full list of publications at `http://sail.usc.edu`
Work funded by NSF SHB program