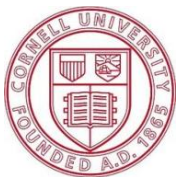

SUBTLE: Overview

(Situation Understanding Bot Through Language and Environment)

Cornell, George Mason, UMass Amherst, UMass Lowell, Penn,
Stanford

Mitch Marcus

University of Pennsylvania



**UMASS
AMHERST**



**STANFORD
UNIVERSITY**



Our Guiding Vision

Standing orders: clear the building: search for any bombs or weapons, find any injured persons and report them and their location so the medical team can find them, and tell any other persons to leave the building.

0:00 Commander: Jr, I want you to search the room you are in first and when you go to leave head for the North end of the building, continue around East and we will meet up at the end of a long hallway.

0:32 Jr: 10-4

3:13 Jr: I have cleared the first room and am proceeding North down a hallway.

3:20 Commander: Mark that room as clear and make certain to check all the rooms in the hall.

3:45 Jr: Room is marked. Proceeding to the next room.

3:56 Commander: Wait. Shouldn't you be going into the hallway?

4:02 Jr: No, I went directly into the next room.

4:05 Commander: Can you show me the map and mark your position.

4:15 Jr: I am displaying the map, I am the yellow dot.



5:23 Commander: Alright I see now, carry on.

Scientific Objective: Central Hypothesis

Effective communications with autonomous bots in real-time situations requires that bots understand not only what is *literally said*, but also what is *intended*, the *implicit meaning*

Example: Exchange between firefighters (Worcester Cold Storage Fire 1999):

P1 Steve, what floor are you on?

P2 I'm on the second floor. Come up that flight of stairs and you go through a freezer chest, and I'm in the back where another freezer chest is. That's where you'll find it.

- P2 (a) *answers question* (b) appears to *give* command
- P1's *implicit* question: How can we find you?
 - Linguists call this the *Question Under Discussion (QUD)*
- *P2 answers the QUD*

Technical Approach: Key Ideas

- We must develop techniques to analyze not only what sentence a speaker used, but also *what that speaker's implicit meaning*.
- The communication system must exploit the *broad context of the environment*.
- The linguistic specification should incorporate *formal models of language*
 - To guarantee *computationally efficient* analysis methods
 - To facilitate study of the *habitability* and effectiveness of the result
- The adequacy of specifications should be determined by *empirical, corpus-based methods*
 - Our research must proceed by collecting example corpora of interactions in increasingly complex (simulated) environments

The SUBTLE team: current co-PIs

Natural Language Processing

- **Mitch Marcus**, *Penn*
- **David Smith**, *UMass Amherst*
- *Aravind Joshi*, *Penn*

Robotics

- **Hadas Kress-Gazit**, *Cornell*
- **Holly Yanco**, *UMass Lowell*
- *George Pappas*, *Penn*

Human Simulation & Graphics

- **Jan Allbeck**, *George Mason*
- **Norm Badler**, *Penn*

Machine Learning:

- **Andrew McCallum**, *UMass Amherst*

Linguistics:

- **Chris Potts**, *Stanford*

Students Actively Involved Year 4

Grad Students

- Cameron Finucane, *Cornell*
- Victoria Schwanda, *UPenn* → *Cornell*
- Weizi Li, *George Mason*
- David Clausen, *Stanford*
- Alex Djalati, *Stanford*
- Sven Lauer, *Stanford*
- Jason Naradowsky, *UMass Amherst*
- Xiaoye “Tiger” Wu, *UMass Amherst*
- Dan Brooks, *UMass Lowell*
- Munjal Desai, *UMass Lowell*
- Eric McCann, *UMass Lowell*
- Abe Shultz, *UMass Lowell*
- Constantine Lignos, *UPenn*
- Qiuye Zhao, *UPenn*
- Alex Shoulson, *UPenn*

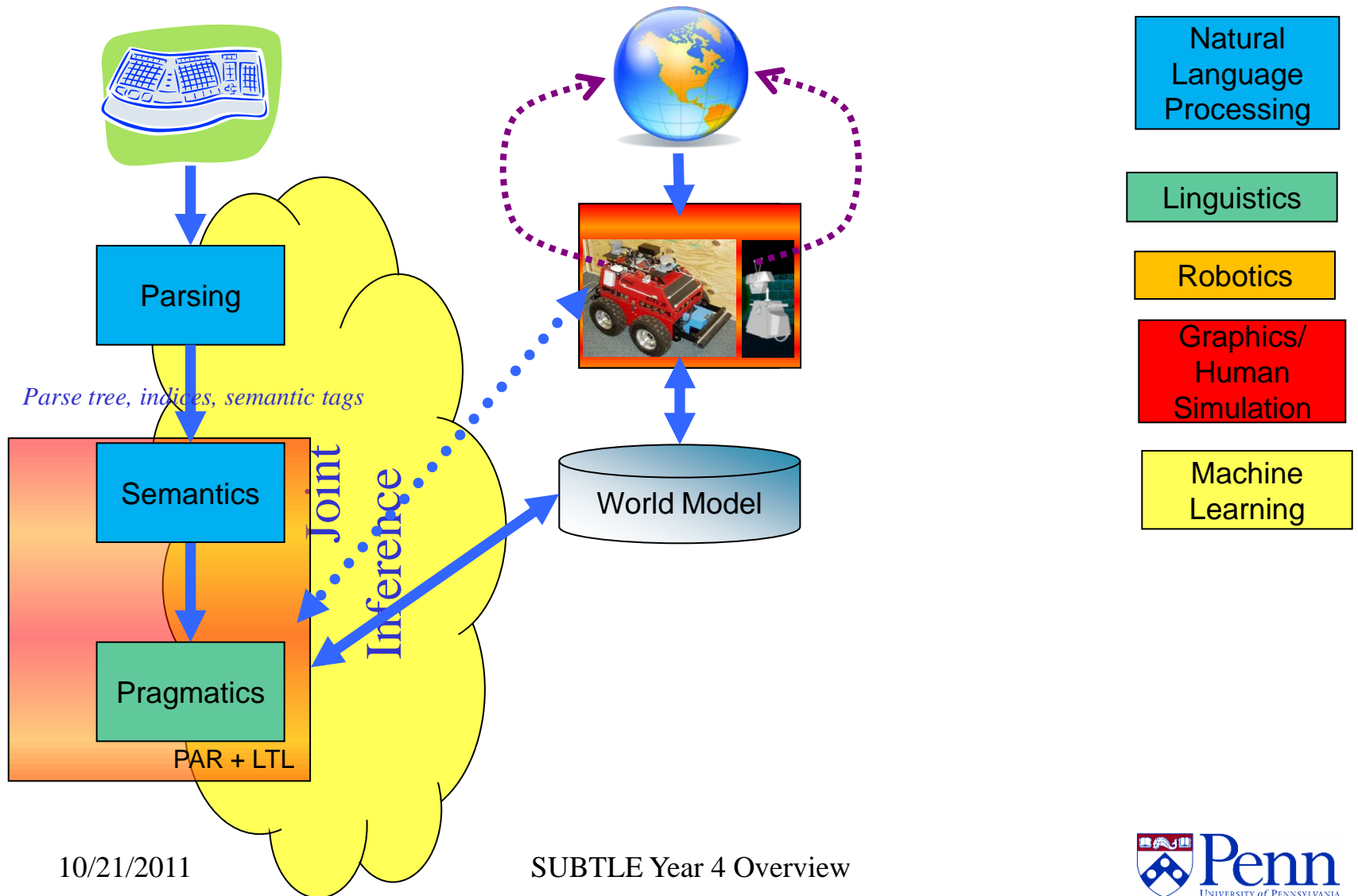
Undergraduates

- Mikahil Medvedev, *UMass Lowell*
- Sean McSheehy, *UMass Lowell*
- [Kenton Lee, UPenn](#)
- [Ian Perera, UPenn](#) → *Rochester*
- Jennifer Bunge, *USMA*

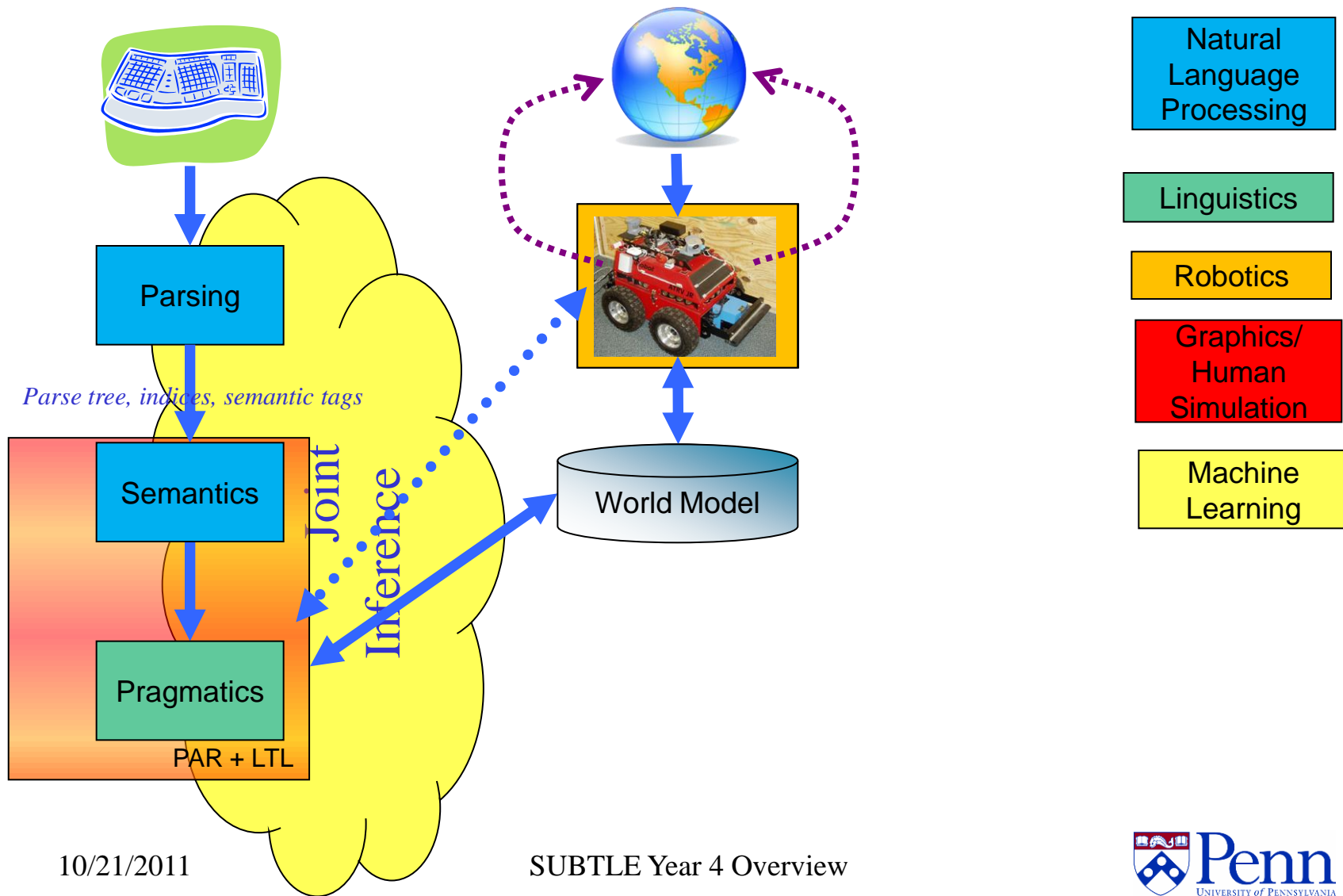
Postdoc

- Ariel Diertani, *UPenn*

Our Target



The Current Reality



Year 1: Sample Accomplishments

- **SUBTLE architecture developed**
- **Key components built and tested**
 - LTL → FSA transducer
 - PAR (Procedural Action Representation) for PragBot
- **“PragBot 1” corpus collection tool built, initial corpus collected, NLP analyzer built**
- **New results: Probabilistic pragmatics meets probabilistic inference**

Year 2: Sample Accomplishments

- **A new formal account of *Grice's maxims*, a foundation of linguistic pragmatics**
 - Embedded and tested in a computation implementation of a rich account of formal linguistic pragmatics
- **Factorie machine inference toolkit under development**
 - Allows efficient machine inference of probabilistic models across very large, complex knowledge bases
- **Initial forays into pragmatics for robot response**



Year 3: Sample Accomplishments

- **Corpus development**

- Large Pragbot 1 corpus collected via Mechanical Turk and annotated
- Pragbot 2 tool developed & debugged.

- **Integration**

- Integrated theory of PAR & LTL developed: MetaPAR
- Initial testbed end-to-end system completed
 - NL Syntax → NL Semantics → MetaPar → LTL → Controller → Graphics Simulation & NL Interaction & Query
- Integration of LTL & UMass Lowell iRobot ATRV Jr: LTLMoP

- **Factorie Dissemination & Initial Use**

Schedule: Year 4 Foci

| | | |
|---------|--|---|
| 8:30am | Introduction | Joe Myers |
| 8:40 | SUBTLE: Overview | Mitch Marcus |
| 9:00 | The SUBTLE NLP Pipeline | Constantine Lignos |
| 9:30 | Semantics and Pragmatics for Pragbot | Ian Perera |
| 9:50 | Break | |
| 10:05 | From language to autonomous robot action: an integrated system (with demonstration) | Dan Brooks, Constantine Lignos & Abe Shultz |
| 11:35 | Understanding reasons for unsynthesizable specifications | Hadas Kress-Gazit |
| | Guaranteeing high-level tasks while exploring | |
| 12:00pm | Lunch | |
| 12:30 | Pragmatic phenomena in shared tasks: studies on the Pragbot I corpus | Chris Potts |
| 1:00 | Pragmatic Inference & Factorie | Tiger Wu |
| 1:20 | Mechanical Turk & the Pragbot II corpus | Kenton Lee |
| 1:40 | Break | |
| 1:55 | Extensions to PAR | Jan Allbeck |
| 2:15 | Event-Based Control of Simulated Human/Robot Interactions using Parameterized Behavior Trees | Norm Badler & Alex Shoulson |
| 2:30 | Wrapup | Mitch Marcus |
| 2:40 | Advisory Committee Meeting | |
| 3:30 | Advisory Committee Initial Feedback | |
| 4:00 | Review Adjourns | |

Third Year Review Recommendations

- ☑ Engage ARL/HRED to develop transition opportunities
- ☑ “Very little about pragmatics and joint inference this year”
- ☑ “Illustrate our improved level of understanding of the model by adding a new level of detail.”
- ☑ “It appears desirable to use MTURK to start collecting data with new scenarios.”
- ☑ “Have a roadmap for addressing some of the brittleness of the system.”
- ☑ “[ARL, NRL, and AFOSR] scientists could pose a similar but somewhat different scenario on which to test the system.”
- ☑ “Look at different kinds of relationships between humans/robots”

Connections to DoD Labs – Year 4

- **Allbeck, Kress-Gazit, Marcus – Visit HRED to investigate technology interaction opportunities 1/5/2010**
- **Marcus–Visit Troy Kelley, HRED, to investigate technology interaction opportunities, 5/5/2010, leading to:**
- **Lignos – 4 weeks at HRED integrating SUBTLE NLP Pipeline components into SS-RICS environment**

Mitch and Joe,
I just wanted to extend a note of gratitude and thanks to both of you and for allowing Constantine Lignos to work with us here at ARL on our robotics program SS-RICS this summer. It has been a very productive summer and we are looking forward to continuing the work over the next several months.
Thanks again,
Troy Kelley
- **Allbeck - Multiple presentations to Night Vision and Electronic Sensor Directorate (NVESD) of the Army at Ft Belvoir. A SOW has been submitted and an avenue for funding planned.**

Connections to DoD Labs II – Year 4

- ***Yanco and McCann*** – Presentation on and demo of multi-touch control the AR Drone, Multi-touch Mission Command Working Group, MITRE Aberdeen, June 2011.
- ***Yanco*** – Visit to APG's SUGV test course for robot test center planning, January 2011.
- ***Yanco*** – Visit to NIST's Standard Test Arenas for Robots for robot test center planning, January 2011.
- ***Yanco*** – National Academies ARLTAB Soldier Systems Panel (May 2011) and ARL Autonomous Systems Review (July 2011), both at APG.