Wubble World

Daniel Hewlett, Shane Hoversten, Wesley Kerr Kerr
Paul Cohen, Yu-Han Chang

University of Southern California
Information Sciences Institute
Introducing Wubble World
Overview

• Goals
• Background
  – Research
  – Gaming
• Games
• Learning Methods
• Current Status and Open Questions
Research Goals

• Wubble World is a virtual environment for learning the semantics of natural language from players

• The meanings of words situated in the virtual environment can be grounded in the objects and actions occurring in it, enabling richer semantics

• The relationship between a player and his/her Wubble is modelled after that of a parent and a language-learning child
Child Language Learning

• Structured language-centric interaction
  – Child: “What's that?”
  – Parent: “That's a dog!”
  – Shared attention, deictic pointing, other factors contribute to the child quickly learning what 'dog' means

• Passive observation of parents and peers
  – Parent 1 holds up a pan, Parent 1 and Parent 2 speak some words to each other, including several instances of 'pan'
  – Over time, child learns what 'pan' means
  – Better than just simple association, because children understand intentions of others (theory of mind)

• Everything in between
Gaming Background

- Children spend hours playing games like Club Penguin, Marian’s World, Neopets
- Much of this time is spent generating language
- *Some* language is about what's going on in the virtual world
  - Could an agent in the same virtual world learn from this language?
Basic Game Model

- Internet games with low system requirements, playable on-demand and for free
- Kid-friendly content
- Each world includes a main “lobby” area where players chat with one another
- Players earn points through playing mini-games
- Points are used to buy customizations
- That's it!
Problems for Language Learning

• Much of the language in such games does not make direct reference to the goings-on in the virtual world
  – Lots of social chatting
    • “I the boss”
    • “No, I the boss!”
    • “You can be co-boss…”

• Mini-games don't offer much help, as they rarely require language or involve other players
Generating Situated Language

• Two Methods
  1) Create mini-games explicitly involving language
     • Language is player-computer
     • 'Blocks Room'
  2) Create multi-player games that require communication about the environment
     • Language is player-player
     • 'Sheep' Game (upcoming)

  'Wictionary' game is an attempt to be both
Blocks Room

• 'Blocks World' in a cozy cottage setting

• Only the Wubble can interact with the room
  – Walk, jump, look around
  – Pick up and put down objects

• The child directs the Wubble using English
“now go in front of the big green cylinder, please”

(ROOT (S
   (ADVP (RB now))
   (VP (VB go)
      (PP (IN in)
         (NP
            (NP (NN front))
            (PP (IN of)
               (NP (DT the) (JJ big) (JJ green) (NN cylinder))))))
   (, ,) (VB please))))
NLU Pipeline II

• The referent of each noun phrase is selected from among the objects in the room
  – With the player's help, if necessary
• Prepositions are resolved to a concrete spatial relation
  – Again, possibly with the player's help
• With the meaning of these words determined, the appropriate action can be executed
Learning in Wubble Room

- Learning is online, human-in-the-loop
- Goal is to learn the meaning of words the child uses
  - Currently: Basic nouns, adjectives, prepositions
  - Future: Verbs, composite objects, and more
- Objects are represented as vectors of discrete features
  - Continuous features such as color or size are ‘binned’
- Learning a word means learning the concept that word denotes
  - probability distributions over feature vectors
The Meanings of Words

• The concept associated with the noun 'cube' would be represented as follows:

  Features
  \- COLOR:
    \- Feature Values
    \- Weights
    \- High-Entropy Weight Distribution
  \- SHAPE:
    \- Feature Values
    \- Weights
    \- Low-Entropy Weight Distribution
  \- SIZE:
    \- Feature Values
    \- Weights
    \- High-Entropy Weight Distribution
Imagination

- Each feature vector has a corresponding probability distribution (derived from the weights on feature values)
- A prototypical object is constructed from sampling these distributions
- All the objects in the room are compared against this prototype, and the most similar is the referent
Feedback

• When uncertain, the Wubble will ask the child to point out the object (and/or location)

• Negative Feedback
  – The Wubble responds by selecting a different object from the room, or querying the player if credit assignment is ambiguous

• Positive Feedback
  – The weights associated with all of the features of the correct referent are strengthened, for each noun and adjective of the corresponding noun phrase
  – Note: No feedback is implicit positive feedback
Avoiding Credit Assignment

• Dialogue:
  – Player: “go to the right of the cube”
  – (Wubble walks to the right of a cylinder)
  – Player: “no”
  – Wubble: “Is this the cube?” (Points to the cylinder)
  – Player: “no”
  – Wubble: “Am I to the right of this?” (Points to the cylinder)
  – Player: “yes”
  – Wubble: “Show me the cube.”
  – (Player points to the cube)
  – (Wubble walks to the right of the cube)
Updating Weights

• Weights are updated so as to minimize regret
  – Method developed by P. Auer, et al.
  – Consider the feature values of the correct object $(v_1, v_2, \ldots, v_n)$, where $v_j$ is a value of feature $f^j$
  – The weights corresponding to each value in each feature vector of the concept is modified according to the following formula (gamma is the reward value):

$$w^j_i(t) = \begin{cases} e^{\gamma}w_i^j(t-1) & \text{if } v_j = i \\ w_i^j(t-1) & \text{otherwise} \end{cases}$$
Qualitative Results

• For most nouns, adjectives, and simple prepositions, a few exposures is enough to produce correct behavior
  – Because player helps the Wubble when it is uncertain, random exploration is minimized

• The assumption of independence between features leads to an inability to precisely learn certain concepts
Quantitative Results

For evaluation of learning performance, see:

Learning in Wubble World
Wesley Kerr, Shane Hoversten, Daniel Hewlett, Paul Cohen, Yu-Han Chang

to appear this July in

Proceedings of the International Conference on Development and Learning (ICDL) 2007
Limitations

• In order for the game to be engaging, the Wubble has to be able to understand (with some help) a good portion of the language the child uses
• This means that the language is constrained to certain types of relatively simple sentences
• Wubble’s knowledge limits complexity of actions
• What if we wanted to learn from open-ended language and activity?
Sentence-Scene Corpus

• By observing players in the course of playing a game and talking about it, a parallel corpus of sentences and scenes can be constructed.

• Ideally
  – 'Sentences' are actually parse trees or logical forms of a wide variety of interesting utterances.
  – 'Scenes' have structured representations at various levels of detail (objects, events, etc.) of a dynamic world.

• With such a corpus, it should be possible to automatically learn associations between the components of the two structures.
Attempt 1: Wictionary

• Players take turns choosing a word or phrase and drawing it on the screen
• Other players must guess the phrase
• Points are awarded to artist and first player who guesses right
Pros

• User-generated content and language means that a relatively simple game can generate a wide variety of structures and labels for them
  – Games like 'Second Life' support content creation in 3D

• Wrong answers might be interesting as well
  – If the right answer was 'duck' and several people guess 'bird', perhaps a relation exists between these words
Cons

- Players stick to short phrases, mostly nouns naming objects
  - 'cow', 'person', 'cactus', etc.
- Though the drawing boards supports some simple dynamics, this feature was rarely utilized
- The need for other players to guess the phrase keeps phrases short
- Ultimately, a different direction was needed to generate more sophisticated situated language
Attempt 2: Sheep

- Team-based multi-player game
Sheep Game

- Team members need to communicate to coordinate their strategy to herd sheep toward their side of the field
  - Voice communication enables real-time language during engaging activity
    - Offline speech recognition for language processing
  - Important game elements require multiple players working together to utilize effectively
  - Player-to-player communication means language is not constrained
Welcome

Wubble World is a fun place to meet new friends and play creative games. Wubbles are curious and intelligent creatures, as you’ll find out when you teach your very own Wubble about the world it lives in. You and your Wubble can compete with other players and win cool accessories.

Enter now to customize your Wubble and begin playing! Or see a preview!