A Machine Translation System from English to American Sign Language

Liwei Zhao, Karin Kipper, William Schuler,
Christian Vogler, Norman Badler, Martha Palmer

Department of Computer and Information Science
University of Pennsylvania, Philadelphia PA 19104-6389 USA
{lwzhao,kipper,schuler,cvogler,badler,mpalmer}@graphics.cis.upenn.edu

Abstract

Research in computational linguistics, computer graphics and autonomous agents has led to the development of increasingly sophisticated communicative agents over the past few years, bringing new perspective to machine translation research. The engineering of language-based smooth, expressive, natural-looking human gestures can give us useful insights into the design principles that have evolved in natural communication between people. In this paper we prototype a machine translation system from English to American Sign Language (ASL), taking into account not only linguistic but also visual and spatial information associated with ASL signs.

1 Introduction

As the third or fourth most widely used language in the United States [23], American Sign Language (ASL) is the primary communication means used by members of the North American deaf community. There is linguistic, psycholinguistic, and neurological evidence in favor of ASL being a fully developed natural language [18]. It is not a derivative of English – it is a complete language with its own unique grammar [11, 28, 30].

While the last ten years have seen an ever increasing development of machine translation systems for translating between major spoken natural languages, translation to and from ASL is virtually ignored by the machine translation community. Yet, ASL translation systems are very important to the deaf. Systems that simply render spoken language as text are inadequate for two reasons: First, many deaf people in the United States have difficulties with reading and writing English; in fact, some do not read above the fourth-grade level. A text-based system would make it impossible for these people to follow and understand a conversation in real-time. Second, if spoken language is rendered as text, all the information on intonation, pitch, and timing is lost, even though this information is important (e.g., the reason why people prefer dubbed movies over subtitled movies). ASL, on the other hand, is capable of conveying this information through the intensity of the signs and facial expressions. As a result, a fully-functional ASL machine translation system would be far superior to a text-based system when it comes to conveying all the nuances of spoken language.

ASL machine translation systems have been neglected largely because of the specialty of ASL as a natural language. Not so long ago ASL was still looked upon as 'merely gestures' – non-linguistic, pantomimic presentations of concrete concepts. For several years some researchers believed that ASL lacked any rigid structure on the sentence-level, which obviously made it very hard to translate any other natural language into ASL. To make things worse, ASL is produced in a modality (or channel) that is greatly different from English: ASL is a signed language; it cannot be spoken; and there is currently no accepted form of written ASL [30]. The earlier commonly-used means of referring to signs in writing is *glosses* notation, whereby signs are represented in their natural order by upper-case words taken from their nearest spoken counterparts. A major drawback of this representation is that it does not show what the translated signs look like.[1] More recent methods use relatively iconic, picture-like symbols to represent the positions and movements of the hands, as well as the facial expressions, but failing to incorporate spatial elements into the representation, this kind of writing system can still cause confusion in complex signs. An ideal approach would use three dimensional (3D) representations of ASL signs being performed, allowing examination from different perspectives, and making accurate understanding and imitation more feasible. This clearly imposes a severe constraint on the target language generation of a machine translation system, however.

Our approach involves two steps: (1) a translation from an input English sentence into an intermediate representation, taking into account aspects of syntactic, grammatical and morphological information; (2) an interpretation of the intermediate representation as a motion representation with a small set of qualitative parameters which can be further converted to a large set of low-level quantitative parameters that actually control the human model to produce ASL signs.

For the intermediate representation, we use *glosses* notation with embedded parameters. To generate the intermediate representation from the input sentence, we need to (i) analyze the word order and figure out which sign order is more appropriate, and (ii) generate the *glosses* and embed parameters indicating grammatical information, such as sentence types, facial expressions, and morphological information. We use a Synchronous Tree Adjoining Grammar (STAG) [21, 20] for mapping this information from English to ASL.

A sign synthesizer is employed for the second step. It assumes that the embedded *glosses* representation is already in correct sign order with appropriately assigned grammatical and morphological parameters. For each sign, it first uses the *gloss* as an index to look up a sign dictionary, which stores the parameterized motion templates for all available ASL signs; then uses the embedded parameters to modify the default parameters defined in the motion template to get the effective parameters. The sign synthesizer employs Parallel Transition

[1] Although the system described in this paper does use *glosses* in its internal intermediate representation, it does not use them in the final output.

Networks (PaT-Nets) [3] to achieve the smooth transitions between signs. PaT-Nets in turn call *Jack Toolkit* and *Jack Visualizer* functions to generate the final animation [8].

We implemented this system and called it `TEAM` (Translation from English to ASL by Machine). The major contributions of this work are:

- To our knowledge, this is the first machine translation system from English to 3D animated ASL, taking into consideration not only linguistic but also visual and spatial information associated with ASL.
- It demonstrates that a machine translation system for ASL that uses full natural language processing and full graphics in real time is feasible.
- Our system is not limited to ASL only. Its flexibility allows it to be easily expanded to other signed languages.

2 Graphics Modeling

In order to output true ASL we need a fully articulated 3D human model. The model should have finely articulated hands, highly expressive arms and body, as well as controllable facial expressions. In addition, we need a fast computational model to procedurally generate a wide range of natural-looking ASL signs.

2.1 Human Model

Our human model (shown in Figure 2.2) has 80 joints with a total of 135 degrees of freedom [8]. The torso is composed of 17 joints in the spine between the waist and the neck. A forward kinematics algorithm is employed to position the torso towards a specified set of joint angles. The movements of the arms are specified through key time and end-effector positions (keypoints). An analytical inverse kinematics algorithm computes shoulder and elbow rotations given a specified keypoint.[2] The hand is finely articulated. We use a left/right-independent library of joint angles to shape the hand into a variety of pre-determined positions. Currently we use MIRALab's Face model [17] for animating facial expressions. Facial expressions play an important role in ASL's grammatical process.

2.2 Effort and Shape

Our approach in generating ASL signs and dynamically changing their motion characteristics is based on recent work [6, 31] on building computational models of a particularly important system called Laban Movement Analysis (LMA). LMA has four major components — Body, Space, Shape, and Effort. The components of LMA that we cover are Effort and Shape [4]. Effort comprises four

[2] An inverse kinematics algorithm computes a set of joint angles that satisfies some constraints, given a desired position and orientation of the end-effector, i.e., the hand. The algorithm we are using is made suitable for an anthropomorphic arm [27].

motion factors: Space (`S`), Weight (`W`), Time (`T`), and Flow (`F`). Each motion factor is a continuum between two extremes: (1) *indulging* in the quality and (2) *fighting* against the quality. These extreme Effort elements are seen as basic, 'irreducible' qualities, which means that they are the smallest units of change in an observed movement. Table 2.2 shows the LMA Effort elements – the extremes for each motion factor. The Shape dimensions in LMA are Horizontal (`H`), Vertical (`V`), Sagittal (`S`), and Flow (`Fl`). The terms used to describe the extreme attitudes towards these dimensions are Spreading and Enclosing, Rising and Sinking, Advancing and Retreating, Opening and Closing, respectively. In general, Shape changes occur in affinities with corresponding Efforts (Table 2 [4]).

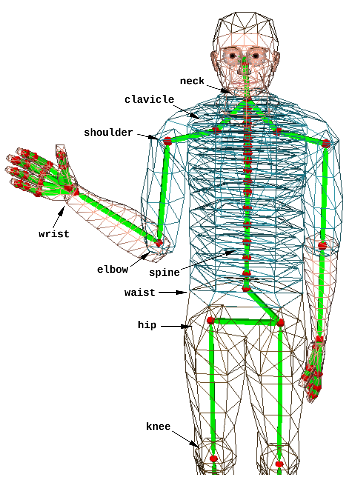

Fig. 1. Human Model

Effort	**Indulging**	**Fighting**
Space	Indirect	Direct
Weight	Light	Strong
Time	Sustained	Sudden
Flow	Free	Bound

Table 1. Effort Elements

Dimension	**Shape**	**Effort**
Vertical	Rising	Weight-Light
	Sinking	Weight-Strong
Horizontal	Enclosing	Space-Direct
	Spreading	Space-Indirect
Sagittal	Retreating	Time-Sudden
	Advancing	Time-Sustained

Table 2. Effort and Shape Affinities

Effort and Shape qualities are expressed using numeric parameters that can vary along distinct scales. Each dimension of Effort and Shape is associated with a scale ranging from -1 to $+1$. Effort parameters are translated into low-level movement parameters which affect the dynamics of the underlying movement, while Shape parameters are used to modify key pose information, which affect the dimensions of space of the underlying movement. For more technical details about Effort and Shape see [6].

3 ASL Linguistics

3.1 ASL Phonology

William Stokoe was the first linguist to recognize and analyze ASL as a language in its own right. In his seminal work [24, 25] a notational system (now called "Stokoe notation") was devised to analyze each ASL sign into three phonological [3] components: *handshape*, *location* (place of articulation), and *movement*. Battison [5] and Frishberg [10] presented evidence that these aspects of the sign, along with *orientation* of the palm of the hand, are internal phonological parameters that are necessary for a complete and efficient description of ASL signs. Changing any of these parameters of a sign may make a new sign and therefore change the meaning of the sign.

Simultaneity and Sequentiality Siple [22] and Kilma and Bellugi [15] have pointed out that the phonological components in ASL exhibit a high degree of simultaneity — they are not produced sequentially. All of the components should be simultaneously present when the signs are produced. However, this does not necessarily mean that sequentiality is not at all incorporated in the phonology of ASL. Linguists [16] proposed that the sequential production of components did indeed need to be represented in ASL phonology, for example, ASL signs move to and from locations.

3D Modeling ASL Phonology Under the analysis above, it is safe to say ASL signs can be sequentially segmented into one or more phonologically significant units, which are simultaneously associated with certain handshapes, locations, and orientations. In order to accurately model an ASL sign, we need to take into consideration all these phonological parameters, as well as the simultaneity and sequentiality characteristics.

Seamless Solutions^{TM} [1] created some interactive 3D avatars that can communicate in ASL, but they did not take into account the phonological information and their avatars basically perform in signed English rather than ASL. SignSynth [12] uses Perl CGIs (Common Gateway Interfaces) to generate 3D animation in VRML format from a phonological specification, but currently only the finger-spelling module is reported to be working. The lack of a good inverse kinematics procedure prevents it from realizing the remaining and more important modules (arm movement, etc.).

Our approach is phonologically-based and more comprehensive. The location and movement for a specific sign are specified by keypoints, which are relative to the human model's shoulders. Keypoints can be either *Goal* or *Via* keypoints. *Goal* keypoints define a general movement path; the hand follows a path which pauses at each *Goal* keypoint. *Via* keypoints direct the motion between *Goal* keypoints without pausing. A fast analytical inverse kinematics algorithm computes the shoulder and elbow rotations, given a keypoint specified by three-dimensional position coordinates and an elbow swivel angel [27]. The base of the wrist acts as the end-effector indicating the keypoints. Wrist orientations (which decide the palm orientations) and handshapes can also be specified.

[3] The terms *phonology* and *phonological* are used to describe ASL sign formation, even though these terms are derived from speech-related phenomena. They are widely used in ASL linguistics research.

3.2 ASL Morphology

ASL differs dramatically from English in the *mechanisms* – inflections and derivations – by which its phonological units are modified.

ASL Inflectional Morphology ASL exhibits a very rich set of inflectional variations [15]. Rather than affix-like sequential additions to signs, inflections in ASL involve superimposed spatial and temporal contrasts affecting the phonological movement. There are quite a few different inflectional processes in ASL. In this paper we focus on inflections for *temporal* aspect, reflecting distinctions such as *frequently*, *slowly*, and *quickly*; inflections for *manner*, such as *carefully*, and *haphazardly*; and inflections for *degree*, such as *very* and *a little bit*. The nuances of meaning expressed by these reflectional processes represent a considerable range of semantic distinctions.

ASL Derivational Morphology In addition to inflectional processes, ASL has a wide variety of methods that expand the phonology by regular systematic changes in the phonological root and result in the formation of related phonological units. At present, we focus on the so-called *paired noun-verbs* [15, 26] which are pairs of noun and verb that share the same phonological characteristics and semantic meaning (for example, *FLY* and *AIRPLANE*, *SIT* and *CHAIR*). To a non-signer the distinction between paired noun-verb may not be readily apparent unless the two are performed consecutively. However, there is a consistent and standardized distinction between them in directionality, manner, frequency, tension, and/or evenness of the underlying movement.

Systematic Changes Underlying Inflectional and Derivational Processes As mentioned, inflectional and derivational morphology can be achieved by systematically changing Effort and Shape parameters. In the following, we show an example of how to embed Effort parameters.

```
TOP-LEVEL
sign / SIGN / SPACE WEIGHT TIME FLOW

SPACE                                    WEIGHT
_[directly]   / _( 0.2, _, _, _)         _[strongly]  / _(_,  0.2, _, _)
_[indirectly] / _(-0.2, _, _, _)         _[lightly]   / _(_, -0.2, _, _)
_             / _( 0.0, _, _, _)         _            / _(_,  0.0, _, _)

TIME                                     FLOW
_[slowly]     / _(_, _, -0.2, _)         _[freely]    / _(_, _, _, -0.2)
_[quickly]    / _(_, _,  0.2, _)         _[boundly]   / _(_, _, _,  0.2)
_             / _(_, _,  0.0, _)         _            / _(_, _, _,  0.0)
```

In this example, *sign slowly*, is translated into embedded *glosses* notation `SIGN(0.0, 0.0, -0.2, 0.0)`. [4] Other inflections as well as derivations can be

[4] The parameter values are set arbitrarily, for example, -0.2 stands for a noticeable change in the Sustained quality.

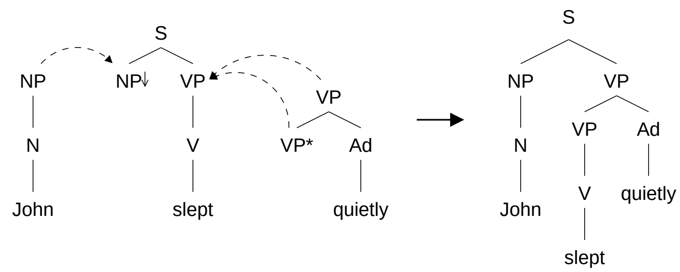

Fig. 2. Substituting and adjoining TAG trees.

done similarly. For example, because the nouns in the paired noun-verbs consistently have a smaller movement that is restrained in dimensions of space, we may assign smaller Shape (or Effort) parameter values to derive the nouns from their associated verbs. Our parameterized representation on the phonology level offers us the advantage that we do not have to list paired nouns and verbs as independent items in the dictionary but can derive one from the other consistently. Our approach also complies with the hypothesis shared by linguistic researchers: that is, the underlying forms of roots may be abstract items that do not occur as surface forms in the language [7].

4 Translating English Sentences

4.1 Synchronous Grammar

We use a Lexicalized Tree Adjoining Grammar based system for translating between English sentences and ASL *glosses*. A Tree-Adjoining Grammar (TAG) is a tree rewriting system [14], the primitive elements of which are *elementary trees* (see Figure 2). In a Lexicalized TAG, these elementary trees are anchored by lexical items, such as nouns and verbs [21]. The elementary trees also have argument positions for the subjects and objects of verbs, adjectives, and other predicates, which constrain the way they can be combined, and which determine the predicate-argument structure of the input sentence. Elementary trees are combined by the operations of *substitution* and *adjunction*, where substituting elementary trees (such as the noun tree for 'John' in Figure 2) are attached at the frontier nodes of other elementary trees (designated with ↓), and adjoining elementary trees (such as the modifier tree for 'quietly') are attached at internal nodes of other elementary trees by removing the part of the host tree below the adjunction site and reattaching it at one of the frontier nodes of the adjoining tree (designated with $*$). A source-language sentence can be parsed using these constrained operations in polynomial time. As the source-language sentence is parsed, a target-language tree can be simultaneously assembled, using Synchronous TAGs [21, 20], by associating one or more target-language elementary trees with each source-language elementary tree, and associating the nodes

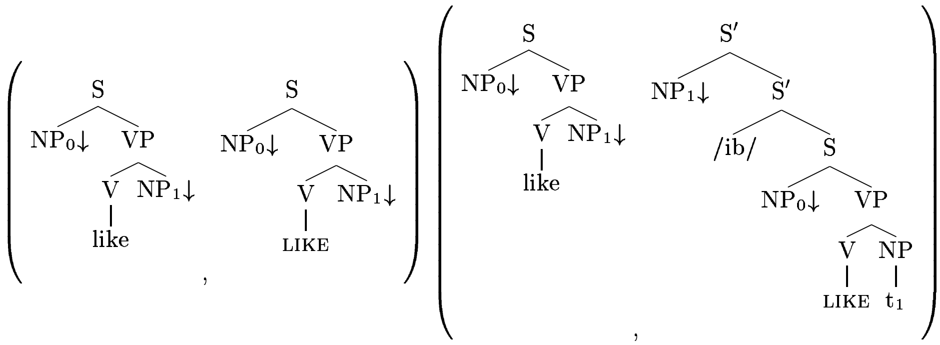

Fig. 3. Untopicalized tree and topicalized tree for pronouns

at which subsequent substitutions or adjunctions can take place. For example, the associated source and target trees in Figure 3, for the English and ASL translations of 'like,' have links (denoted by subscripts) between the designated sites for substitutions of noun trees (like 'John' in the previous example), which will function as the subjects in each sentence. Recognizing a complete parse on the source (English) side therefore means building a complete parse on the target (ASL) side. Synchronous TAGs have been used for machine translation between spoken languages [2] but this is the first application to a signed language.

The input sentence brings with it grammatical information such as sentence types and morphological marks such as tense. This information is expressed in ASL through non-manual signals which have to be incorporated into the target derivation tree.

4.2 Grammar patterns

Although ASL seems to have a freer word order than English, Fisher [9] claims that the underlying structure of ASL is based on subject-verb-object (SVO) order and when other patterns are used, they need to be marked by 'intonational breaks' (e.g., pauses, head tilts, raising of eyebrows). Topicalized ordering is preferred when the sentence has pronouns or when the discourse referents are present.

Intransitive and transitive verbs can be represented either in SVO or in topicalized order; the choice of order depends heavily on the use of nouns and/or pronouns. Translations (1) and (2) show the sentences in English and the respective ASL *glosses* with the addition of an *intonation break* (/ib/) which is mapped to the non-manual signs for topicalization in the animation system. Figure 3 shows the tree mappings for these two pairs of sentences.

(1) *John likes the girl*
JOHN LIKE GIRL

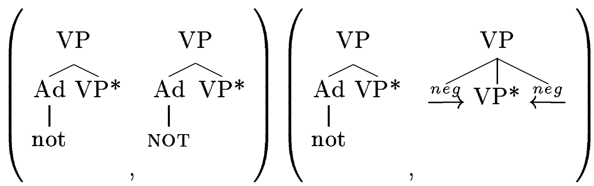

Fig. 4. Synchronous tree pairs for manual and nonmanual negation

(2) *I like the girl*
GIRL */intonation_break/* I LIKE

Predicative constructions in English which have nouns and adjectives as descriptors are mapped to ASL sentences without verbs. The translation in (3) is an example of this construction which also allows for topicalization.

(3) *They are doctors*
DOCTOR */intonation_break/* HE SHE

Negation in ASL may include the *gloss* NOT or simply movement of the head side-to-side to express the idea of negation. It is usually accompanied by a nonneutral facial expression, such as a frown. We use the top line notation to specify the scope of the non-manual signs to be performed by the animation system. The mapping between English and ASL shown in Figure 4 includes the beginning ($\xrightarrow{neg}$) and ending ($\xleftarrow{neg}$) marks for these signs in the ASL tree. In a sentence such as (4), the non-manual signs expressing negation are performed throughout the VP, "NOT BUY DOG."

(4) *The woman is not buying a dog*
WOMAN $\overline{\text{NOT BUY DOG}}^{\text{neg}}$

A declarative sentence, a yes/no question and an imperative sentence may all map to the same *glosses* in ASL. The translations in (5), (6), and (7) show how these three sentence types are distinguished by the way they are presented by the signer through the use of non-manual signals. The declarative sentence in (5) is delivered at a normal rate and with a neutral face; the yes/no question in (6) is delivered with a question sign at the end, which means it requires a head tilt and eyebrows raised to convey the question idea and also the arms do not go back to a neutral position immediately after; and the imperative sentence in (7) is delivered with lower eyebrows and eyes slightly squinted, and the verb is signed with more force, which we map using the Effort parameters. PAST is a sign gesture with the hand above the shoulder.

(5) *I left the cat in the house*
PAST */intonation_break/* LEAVE CAT IN HOUSE

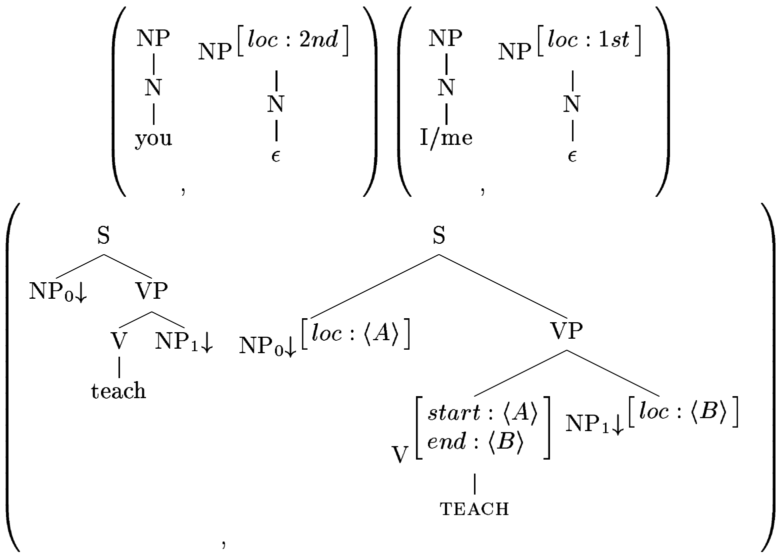

Fig. 5. Synchronous tree pairs for "I teach you," and "you teach me"

(6) *Did you leave the cat in the house?*

PAST */intonation_break/* $\overline{\text{LEAVE CAT IN HOUSE}}^{\,q}$

(7) *Leave the cat in the house*
LEAVE(0,0.5,0.2,0.2) CAT IN HOUSE

There are different classes of verbs in ASL. One distinction is between *nondirectional* and *directional* verbs. In the first case, the sign is always the same and cannot be changed without altering its meaning. If a nondirectional verb is accompanied by pronouns, these need to be mapped explicitly and must be present as *glosses* in ASL. Examples of these verbs include KNOW, FORGET, EAT, and TRUST. For directional verbs, such as ASK, GIVE, WAIT, and TEACH, the start and end points of the verb sign depend on the pronouns used (1st, 2nd or 3rd person). For example, TEACH has a different orientation for *You teach me* (8) than for *I teach you* (9). In this case, we map the pronouns to ϵ (empty arguments) and have features identifying their locations in space which are used as start and end points for the verb sign. Other classes of verbs also exist, such as location verbs, where only the direct object is referred to; we treat these the same as directional verbs but with the start point in neutral position.

(8) *I teach you*
$_1\text{TEACH}_2$

(9) *You teach me*
$_2\text{TEACH}_1$

For multi-directional verbs such as *teach*, the translation must match the start- and end-points of the sign for the verb with the locations of the signs for the subject and object. This behavior, which closely resembles subject-verb agreement in spoken languages such as English, is modeled using essentially the same mechanism used for subject-verb agreement in existing grammars [13]. We handle this using features in the paired grammar which coindex the location of the verb's subject and object signs with features on the verb sign, which are interpreted as parameters for the start and end points of the sign in the parameterized gesture system (Figure 5).

Fig. 6. An ASL Animation Example: "RECENTLY I-SICK NOW I-WELL"

Incorporation of number, size and shape, manner, and location, among others, occurs frequently in ASL. These phenomena present problems for translation, since they are frequently dependent on the signer's understanding of the sentence.[5] We use the Effort-Shape parameters discussed in (2.2) to handle inflections for *temporal aspect*, *manner*, and *degree*. The example translation in (10) shows the mapping of the adverb *slowly* to the appropriate parameters that modify the way the sign OPEN is performed.

(10) *John opened the door slowly*
JOHN OPEN(0,0,-0.2,0) DOOR

Conditional sentences are expressed in English by using the lexical item *if*. In ASL one way to express a conditional sentence is by marking the condition

[5] For example, in the signed version of the English sentence *He picked up the dead fish and threw it in the trash* it is likely that the *gloss* FISH will be signed as far as possible from the signer's body.

as a question and the consequent as a statement. Examples of this can be seen in (11) and (12).

(11) *If it rains tomorrow, I will not go to the beach*
$\overline{\text{TOMORROW RAIN}}^{\,q}$ $\overline{\text{ME NOT GO BEACH}}^{\,neg}$

(12) *If it rains tomorrow, are you going to the beach?*
$\overline{\text{TOMORROW RAIN}}^{\,q}$ $\overline{\text{YOU GO BEACH}}^{\,q}$

5 Sign Synthesizer

In most cases, the transitions between the signs should be smooth. A simple and straightforward approach for the smoothness is to have the beginning and ending of every sign performed in the same standard posture. While this approach offers smooth continuous transitions, beginning and ending each sign in the same 'neutral' position is very unnatural. An awkward, computationally expensive approach is to define transitions between every pair of possible motions. NYU's Improv project [19] uses a technique called *motion blending* to automatically generate smooth transitions between isolated motions. This approach succeeds in avoiding returning to a required 'neutral' pose, but it does not necessarily guarantee natural and rational transitions.

We are using PaT-Nets (Parallel Transition Networks) to solve the motion blending problems. A PaT-Net is a simultaneously executing finite state automata in which the nodes are associated with actions and connections between the nodes are associated with transition conditions. PaT-Nets make it very easy to wait on completion of actions before moving onto the next action, to execute actions in parallel or in sequence, and to dynamically extend the action structure by invoking other PaT-Nets from nodes of the current one [3].

To demonstrate the power of PaT-Nets and the integration of our `TEAM` system, we create an ASL animation example with English input "recently I was sick but now I am well" (Figure 6). The animation was generated in real time.

6 Conclusion and Future Work

We have described a prototype machine translation system from English to American Sign Language, taking into account not only linguistic but also visual and spatial information associated with ASL signs. Currently we focus on translation from English to ASL, but translating the other way around from ASL to English is an even more interesting challenge. If we can create an ASL recognizer [29] and parser as we have for English, we would be able to translate signed ASL to spoken English sentences and back again to allow real-time interaction on the Internet.

7 Acknowledgments

This research is partially supported by U.S. Air Force F41624-97-D-5002, Office of Naval Research K-5-55043/3916-1552793, DURIP N0001497-1-0396, and AASERTs N00014-97-1-0603 and N0014-97-1-0605, DARPA SB-MDA-97-2951001, NSF IRI95-04372, SBR-8900230, IIS99-00297, and EIA-98-09209, Army Research Office ASSERT DAA 655-981-0147, NASA NRA NAG 5-3990, AFOSR F49620-98-1-0434 and Engineering Animation Inc. (EAI).

References

1. *Seamless Solution Inc.*, 1998. http://www.seamless-solutions.com.
2. Anne Abeillé, Yves Schabes, and Aravind K. Joshi. Using lexicalized tree adjoining grammars for machine translation. In *Proc. of the 13th International Conference on Computational Linguistics (COLING '90)*, Helsinki, Finland, Aug 1990.
3. Norman Badler, Martha Palmer, and Rama Bindiganavale. Animation control for real-time virtual humans. *Communications of the ACM*, 42(8):65–73, 1999.
4. Irmgard Bartenieff and Dori Lewis. *Coping With the Environment.* Gordon and Breach Science Publishers, New York, 1980.
5. Robin Battison. Phonological deletion in American Sign Language. *Sign Language Studies*, 5:1–19, 1974.
6. Diane Chi, Monica Costa, Liwei Zhao, and Norman Badler. The EMOTE model for Effort and Shape. In *Proceedings of SIGGRAPH*, New Orleans, LA, 2000. ACM Computer Graphics Annual Conference.
7. Noam Chomsky and Morris Halle. *The Sound Pattern of English.* Harper, New York, NY, 1968.
8. Engineering Animation Inc. *Jack Toolkit 2.2 Reference Guide*, 1999.
9. Susan Fisher. Influences in word order change in ASL. In C. Li, editor, *Word order and word order change*, pages 1–25. Univ. of Texas Press, Austin, TX, 1975.
10. Nancy Frishberg. Arbitrariness and iconicity: Historical change in American Sign Language. *Language*, 51(3):696–719, 1975.
11. Victoria Fromkin. Sign language: Evidence for language universals and the linguistic capacity of the human brain. *Sign Language Studies*, 59:115–127, 1988.
12. Angus Grieve-Smith. Sign synthesis and sign phonology. In *Proceedings of the First High Desert Student Conference in Linguistics*, New Mexico, 1998.
13. XTAG Research Group. A lexicalized tree adjoining grammar for English. Technical report, University of Pennsylvania, 1998.
14. Aravind Joshi. An introduction to tree adjoining grammars. In A. Manaster-Ramer, editor, *Mathematics of Language.* John Benjamins, Amsterdam, 1987.
15. Edward Klima and Ursula Bellugi. *The Signs of Language.* Harvard University Press, 1979.
16. S. K. Liddell and R. E. Johnson. American Sign Language: The phonological base. *Sign Language Studies*, 64:195–277, 1989.
17. Nadia Magnenat-Thalmann, Prem Kalra, and Marc Escher. Face to virtual face. *Proceedings of the IEEE*, 86(5), 1998.
18. Lynn Messing and Ruth Campbell. *Gesture, Speech and Sign.* Oxford University Press, 1999.
19. Ken Perlin. Real time responsive animation with personality. *IEEE Transactions on Visualization and Computer Graphics*, 1(1):5–15, March 1995.

20. Stuart Shieber. Restricting the weak-generative capability of synchronous tree adjoining grammars. *Computational Intelligence*, 10(4), 1994.
21. Stuart Shieber and Yves Schabes. Synchronous tree adjoining grammars. In *Proceedings of the 13th International Conference on Computational Linguistics (COLING '90)*, Helsinki, Finland, August 1990.
22. Patricia Siple. *Understanding Language Through Sign Language Research.* Academic Press, 1976.
23. Martin Sternberg. *The American Sign Language Dictionary.* Multicom, 1996.
24. William Stokoe. *Sign language structure. Studies in Linguistics: Occasional Paper 8.* Buffalo, New York, 1960.
25. William Stokoe, Dorothy Casterline, and Carl Croneberg. *A Dictionary of American Sign Language on Linguistic Principles.* Gallaudet College Press, Washington, DC, 1965, 1976 (2nd ed.).
26. Ted Supalla and Elissa Newport. How many seats in a chair? The derivation of nouns and verbs in American Sign Language. In P. Siple, editor, *Understanding Language through Sign Language Research.* Academic Press, New York, NY, 1978.
27. Deepak Tolani. *Inverse Kinematics Methods for Human Modeling and Simulation.* PhD thesis, University of Pennsylvania, 1998.
28. Clayton Valli and Ceil Lucas. *Linguistics of ASL: An Introduction.* Gallaudet University Press, 1995.
29. C. Vogler and D. Metaxas. A framework for recognizing the simultaneous aspects of American Sign Language. To appear in Computer Vision and Image Understanding.
30. Sherman Wilcox and Phillys Wilcox. *Learning to See.* Gallaudet University Press, Washington, DC, 1996.
31. Liwei Zhao, Monica Costa, and Norman Badler. Interpreting movement manner. In *Computer Animation 2000.* IEEE Computer Graphics Society, May 2000.