Case-Based Reasoning for Planning Turn-Taking Strategy with a Therapeutic Robot Playmate

Hae Won Park and Ayanna M. Howard

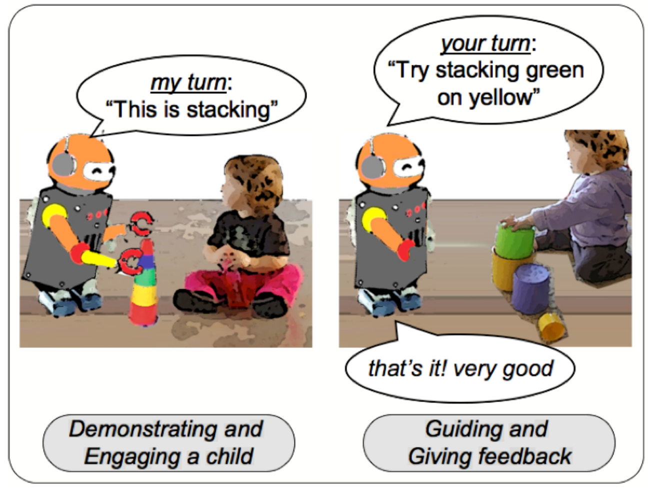
***Abstract*—In this paper, we focus on robot intelligence to generate turn-taking strategies in response to human play actions. This work builds on our previous work on play behavior recognition, and expands it to the child-robot therapeutic domain where the robot must understand and learn the play of a child and take turns manipulating the toys. The main contribution of this work is a novel attempt in applying Case-Based Reasoning (CBR) for planning human-robot turn-taking strategies. By comparing the child's play in the current scene to some past play cases stored in memory, we retrieve the best solution and adapt it to the set of toys that are available for the play scenario, bypassing a long complicated decision process. In order to ensure real-time performance, a low dimension scale invariant shape descriptor is proposed for shape matching. Turn-taking CBR (ttCBR) system is then evaluated for stacking and inserting tasks with four subjects, by comparing the decision made by the system and the actual choice of the humans.**

I. INTRODUCTION

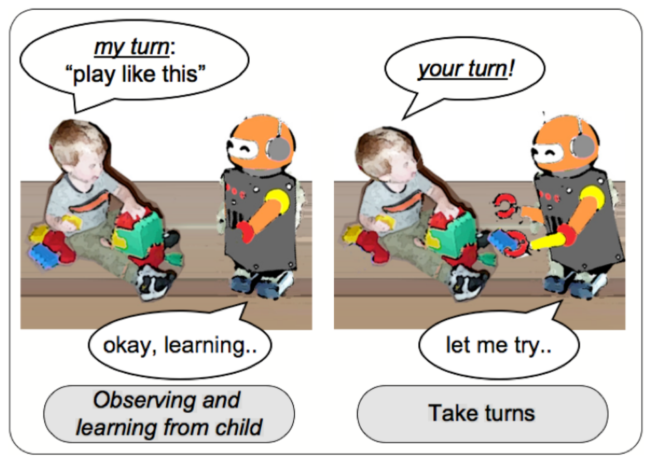
MY turn! Your turn! Turn-taking is a crucial part of children's everyday play activities. Through turn-taking, young children learn basic interaction skills, collaboration, patience, and build understanding of others [1]. However, for children with developmental disabilities, the turn-taking concept can be quite difficult to comprehend. For such children, uniform and repetitive exposure to interaction play is important [2]. In real life, parents may lack knowledge, time, or patience to provide such play. Trained therapists, on the other hand, can provide quality sessions, but are burdened by cost and are often time limited.

In our previous work [3], a method to understand a child's play behavior was introduced as a first step to engage children in play. Play behaviors were modeled by sequencing low-level play primitives, which provides versatility to understand any kind of motions. Basic motions that form play manipulations described in Baranek's list [4] were studied through YouTube videos. Afterwards, fourteen play primitives were defined and trained by Hidden Markov Models (HMMs). The toys in the scene were detected and tracked by the histogram back-projection method using illumination invariant 2D hue-saturation histogram model.

In recent years, many promising works on robot playmates for children have been developed. They give commands to children, react to senses, express emotions,



(a) Therapy turn-taking play led by the Robot



(b) Turn-taking play led by the Child

Fig. 1. Collaborative turn-taking play in childhood is an important source for learning social behaviors. Therapeutic robot playmates have the reasoning ability to plan a turn-taking strategy with CBR framework. Possible playmate roles are shown in (a) and (b).

and emit sound or verbal responses [5-9]. While these research efforts focus on engaging children in interactive play, we propose a reasoning system that goes beyond basic interaction, encompassing the social aspect of turn-taking. Turn-taking play requires a reasoning process that interprets the given problem and deduces a solution. Case-Based Reasoning (CBR) is a concept that solves new problems based on the solutions of similar past problems [10].

The main contribution of this work is a novel attempt in applying CBR for synthesizing turn-taking strategies in child-robot interaction. By comparing the child's play in the

Manuscript received April 9, 2010. This work was supported in part by the National Science Foundation under Award Number IIS-0705130.

Hae Won Park and Ayanna Howard are with the School of Electrical and Computer Engineering at Georgia Institute of Technology, Atlanta, GA, USA. (e-mail: haewon.park@gatech.edu).

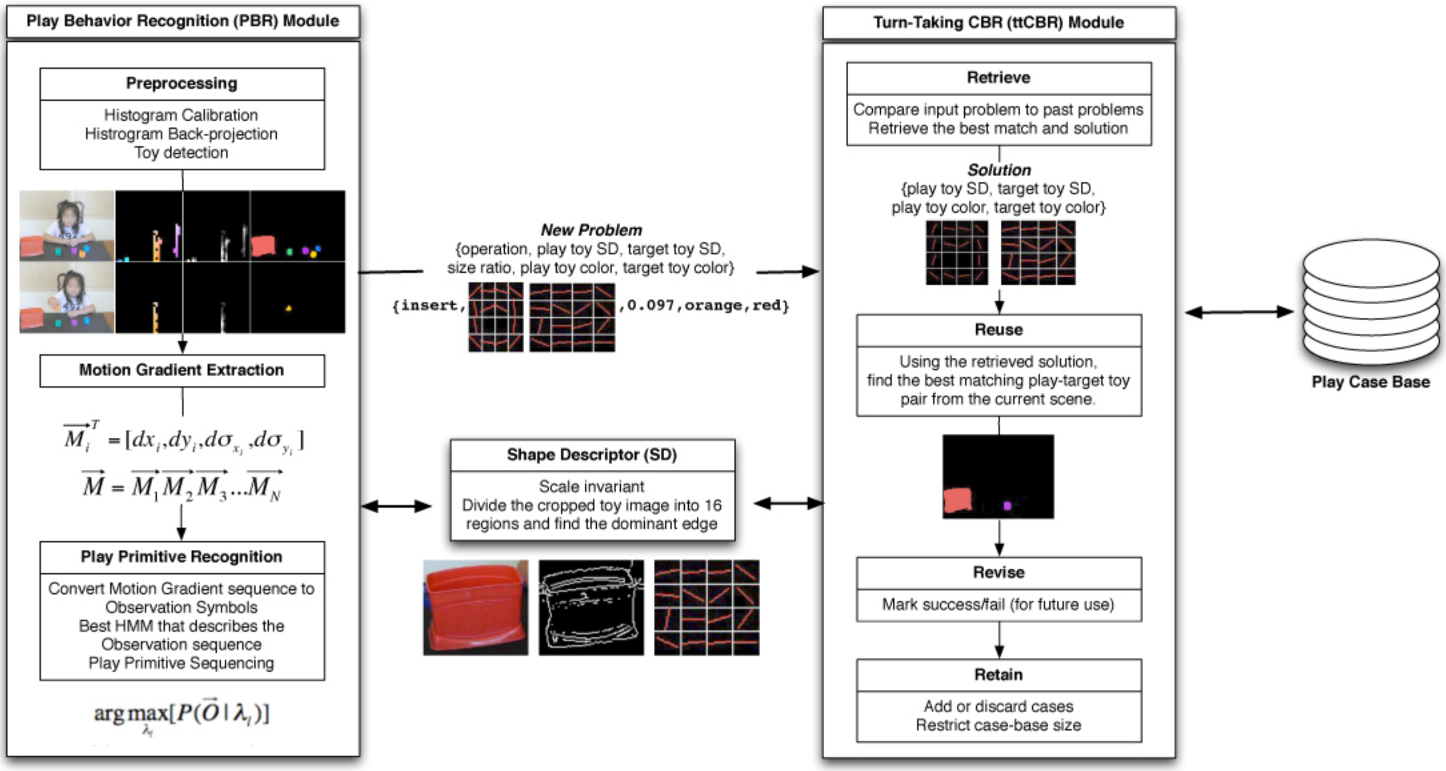


Fig. 2. The overall system architecture. The Play Behavior Recognition (PRB) module from the previous study [3] detects play behavior and toys in the scene, and outputs a new problem. The turn-taking Case-Based Reasoning (ttCBR) module compares the problem to existing play cases and retrieves the solution of the best matching case. The solution is modified during the reuse stage in order to adapt to the current scene, and is added to or deleted from the Play Case-Base.

current scene to some past play cases stored in memory, we retrieve the best set of solutions that matches the play scenario, bypassing a long complicated decision process.

The remainder of this paper is organized as follows. Section II gives a literature review of the existing approaches to the human-robot interaction problem, with specific discussion on real-time CBR systems. Section III presents the general architecture of our system, discussing issues such as case structures, and retrieve/reuse/revise/retain steps of the CBR framework. Low dimensional scale-invariant shape descriptors, the essential component of real-time performance, will also be explained in details. Results are discussed in Section IV, emphasizing the advantage of our turn-taking CBR framework. Section V gives concluding remarks and directions for future work.

II. Related Work

A turn-taking social behavior begins by imitating others. By studying children's turn-taking plays via videos, we analyzed how children behave when their turn comes and ends. Most of the videos have shown that after an adult or another child partner takes their turn, a child imitates the exact same actions being demonstrated [11]. While typically developing children possess the ability to imitate others from birth, children with a developmental disorder, such as autism, demonstrate significant difficulty in object imitation and motor imitation [12-14]. In children with autism, imitation skills are thought to be closely related to early language and social abilities [15]. Dawson *et. al*'s work [16] shows that repetitive training and exposure to imitative turn-taking toy play for autistic children elongates the length of an eye contact with their play partner. Therefore, a therapeutic robot playmate with turn-taking strategy has great potential in being applied to numerous play therapies. If the robot has knowledge of a certain therapy play in advance, and has the ability to evaluate the child's behavior, it can guide and assist the child in achieving the objective of the intended play (Fig. 1. (a)). On the other hand, the robot can also participate in a child-led play, where the robot imitates a play created by the child (Fig. 1. (b)).

In the field of robotics, there has been some research that demonstrates taking turns between a human and a robot. While there have been a number of works focused on conversational turn-taking [17-18], there have not been many action or behavior related turn-taking methods explored in the field of human-robot interaction. As a social mediator, the child-like humanoid Kaspar has shown potential in encouraging autistic children to participate in an imitation play [5]. In this work, the child and the therapist take turns imitating Kaspar's expressions. Some children, after observing the robot play with a tambourine, mimic the

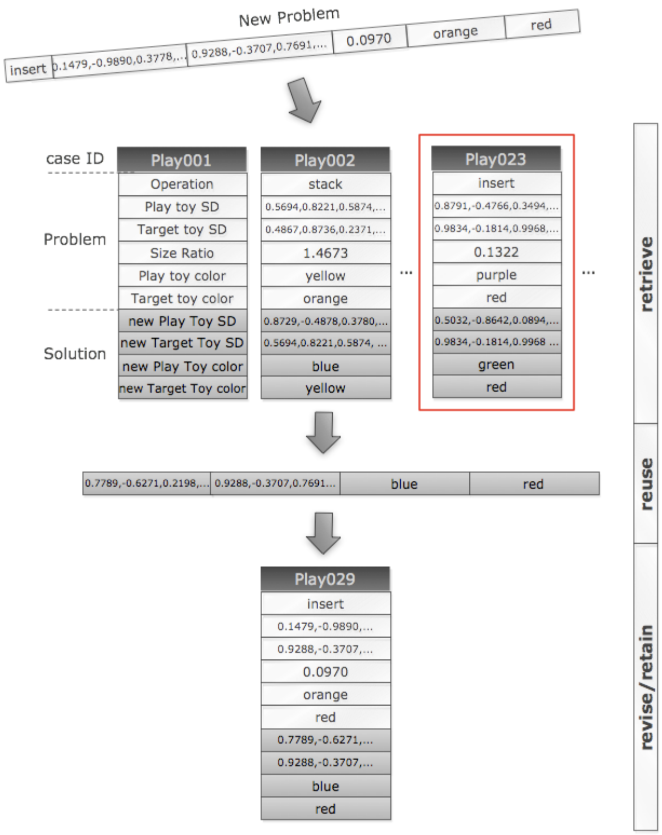


Fig. 3. Turn-taking CBR (ttCBR) system. Play case structure consists of three components: *case ID*, *problem description*, and *solution*. Play case-base is maintained by four stages of Case-Based Reasoning.

action. The result of the study promises that repeated exposure to therapy robots enhances the social skills of autistic children. Compared to this work, our goal is to develop a robotic system that could observe the child's play and produce its own turn-taking strategy. In our turn-taking therapy play, we use toys to provide a strict meaning of a *turn*. Taking a turn means deciding which toys to play with, and this step involves imitation and reasoning. In order to mimic the partner's play, one has to plan for the appropriate set of toys, perhaps deducing from past experience. This is the main reason we take a novel attempt to develop a turn-taking Case-Based Reasoning (ttCBR) strategy for a robot playmate.

III. Approach

Before starting this section, we will introduce some important terminologies in Case-Based Reasoning that is used throughout this paper. The overall architecture of the system is introduced in Fig. 2, and the turn-taking Case - Based Reasoning (ttCBR) module is highlighted in Fig. 3. In Fig. 3, a new *problem* is introduced to the CBR system. The problems, indicated by lightly shaded regions, describe what play has been executed in each turn. For instance, the new problem in Fig. 3 reads, an orange toy was inserted into a red toy, and the size ratio between the two toys is 0.0970. The vectors in the second and third columns define the shapes of the toys. The medium shaded regions define a *solution*. The solution is a set of toys chosen in the succeeding turn. Lastly, the dark shaded region is a *case ID*. The case ID, problem, and solution together define a *play case*. The structure of a play case is illustrated under the case example Play001 in Fig. 3, and also will be explained in details in Section III-B. During the initial training session, play cases are gathered to build a *play case-base*.

The Case-Based Reasoning has several advantages in human-robot interaction tasks. The main advantage is that it can bypass complicated computations during the decision-making process. In this application, we simplify the computation into identifying a play set from the current scene that is most similar to the solution retrieved from the case-base. Another advantage is that it is relatively easy to set up and maintain a knowledge base. CBR has a firm four-stage structure, retrieve-reuse-revise-retain, that is necessary and sufficient to manage and maintain cases. Every case can be easily added to, or deleted from the case-base through these steps.

However, most CBR applications are not real-time systems. The majority of CBR systems utilize their reasoning to aim for high accuracy with less importance on speed. This approach is reasonable for some fields, such as medical diagnosis and travel agents, because given the volume of possible solutions, it is critical to provide a detailed and accurate solution regardless of the time. In our application, child-robot interaction, it is important that the decision is made in real-time in order to ensure the proper response at the appropriate instance. Therefore, robust real-time performance is the first goal for our turn-taking CBR system. The key contributions for ensuring a real-time system are the computation time and size reduction of each play case. We have developed a computationally inexpensive way to match shapes using a scale and intensity invariant, low dimensional shape descriptor.

A. Architecture

The overall architecture of the system is depicted in Fig. 2. The play behavior recognition (PBR) module is responsible for recognizing the subject's play behavior including stacking and inserting actions [3]. The PBR module combines six play components to form a problem: play type (insert/stack), shapes and colors of the play/target toys, and the size ratio between the two toys. Once defined, this problem becomes an input to the turn-taking CBR (ttCBR) module. The ttCBR outputs a new set of toys to be played in the following turn. The stages and procedures will be explained further in Section III-D.

B. Case Structure

Each play case consists of three parts: case ID, problem description, and solution. The case IDs are given to the cases stored in the case-base, the problem description illustrates the play scenario, and the solution is what the past problem deduced as succeeding play objects.

<u>Problem Descriptor</u>

1) *Operation*: The play operation is distinguished by an interaction between the play toy and the target. In this paper, stack and insert plays were evaluated through ttCBR system.

2) *Play toy and target toy shape descriptor*: The Shape Descriptor (SD) unit computes a 32 dimensional vector that describes the scale and intensity invariant shape of the toy.

3) *Play and target toy colo*r: The color of the toy is the least significant property, but when strong multiple candidate solutions are found, it may be used as the last similarity measure.

4) *Size ratio*: The play toy size relative to the target defines the size ratio. Currently, the size ratio roughly depends on the pixel size.

<u>Solution Descriptor</u>

1) *New play toy and target toy shape descriptor*: The 32 dimensional shape descriptors of the toys for the succeeding turn.

2) *New play and target toy color*: The least significant property in finding a new solution, but when strong multiple matching toys are found, it is used as the last similarity measure.

C. *Shape Descriptor*

When playing with nesting cups, or stacking rings, the robot has to possess shape recognition ability. These kinds of toys have multiple parts with different size and color, but share the common outlining shape. The dimension of the shape descriptor has a direct impact on the speed of the retrieval step, which is crucial for real-time performance. Therefore, the goal of our shape matching algorithm is to design a descriptor that is computationally inexpensive, has low dimension vector, and still has robust scale/intensity invariance. Popular object recognition techniques like Scale-Invariant Feature Transform [19] or Speeded Up Robust Features [20] extracts interesting features to provide feature description of the objects. However, since many of the toys children play with lack distinctive features due to their smooth surface and edges, these feature-based object description approaches are inappropriate for our work. Also, the above algorithms are computationally expensive and thus difficult to implement in real-time. The scale-invariant Template Matching [21] may be executed faster, but it is still insufficient for real-time CBR systems.

The shape descriptor we propose in this paper computes dominant edge angles that describe the shape. It is like a rough sketch of the object with linear lines. After the toys are detected in the PBR module, each toy scene is cropped with a minimal margin. The cropped image is then converted into grayscale, and the edge map is computed. The edge map is then divided into 16 regions. For each region, the dominant linear edge angle is computed by the linear least squares fitting. The process is shown in Fig. 4. The angle of the lines in each region defines the shape of the toy.

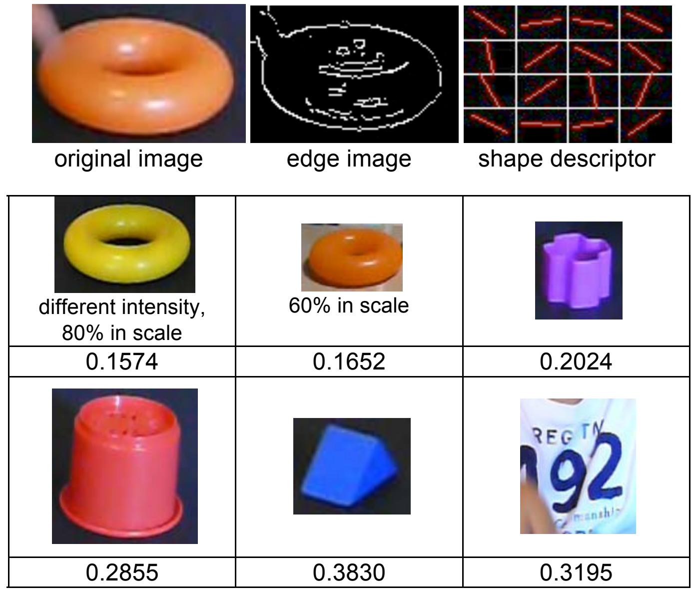


Fig. 4. Example of Shape Matching. A toy is converted into the shape descriptor (top), and is being matched to various objects in the scene. The numbers under each cropped scene represent the distance between the two shape descriptors.

In order to evaluate the proposed shape descriptor, we have collected 40 cropped images that consist of 30 toys and 10 random backgrounds from the scene. If the distance between the shape descriptors is within a threshold, the two toys are classified as the same shape. The average time it took for comparing a toy within this set was 0.62 ms. The maximum average successful recognition rate was 70%. The recognition rate was computed by dividing the sum of true positives and true negatives by the total number of test images. From Fig. 4, we can see that the descriptor performs quite well for different scales.

D. *Turn-Taking Case-Based Reasoning (ttCBR)*

1) *Retrieve*

The retrieval stage is where we compare the problem description with the cases in the case-base, and find the best matching past problem and its solution.

For real-time retrieval, we have created a simple yet effective scale-invariant shape descriptor which has low dimension, reasonable recognition rate, and minimal computation time. For similarity measurements, we have used different methods for each category. First, the operation type of the cases has to match that of the requested problem. The Euclidean distance is used to compute a similarity between the shape descriptors and the size ratio. Given the fact that some plays (e.g., nesting/stacking) require strict shape matching, some depend on the size ratio (e.g., insert), and others rely on both (e.g., block stacking), we have run experiments by assigning different weights on the shape similarity and the size ratio. In various play settings, the

(a) (b) (c)

	Play (a)	Play (b)	Play (c)
Problem (Play Scene)	Subject *stacks* *blue cup* on *orange cup*. Size ratio 0.85	Subject *inserts* *blue block* into *red bin*. Size ratio 0.11	Subject *inserts* *orange block* into *red bin*. Size ratio 0.14
Retrieve			
ttCBR retrieved Problem	stack blue cup on orange cup	insert blue block into red bin	insert blue block into red bin
ttCBR retrieved Solution	stack red cup on blue cup	insert purple block into red bin	insert purple block into red bin
Reuse			
ttCBR adapted Solution	stack {red cup} on blue cup	insert {orange,purple , green block} into red bin	insert {purple block} into red bin
compared to			
Subject's Solution	stack red cup on blue cup	insert orange block into red bin	insert blue block into red bin

Fig. 5. Play (a), (b), and (c) are being evaluated with ttCBR. After each given problem is analyzed, ttCBR retrieves the best matching problem and its solution from the play case-base. The ttCBR adapts the solution to the current play scene and compares it to the subject's solution.

retrieval performed most reliably when the two were weighted equally. The toy colors played a small role when there were multiple closely scored candidates.

2) *Reuse*

The reuse stage adapts retrieved solution to the current scene. This means finding a set of play-target toys that matches the old solution the best. The remaining toys in the play scene are cropped into images in the PBR module, then are converted into shape descriptors in the SD unit. Finally, the resulting shape descriptors are compared to the retrieved solution. Again, the shape descriptor and the size ratio contribute equally, and the toy colors play a least significant role when there are multiple high scored candidates. This process can be easily reduced when target or play toy remains the same, or play toy becomes the target.

3) *Revise/Retain*

The revise stage is to be added once the system is integrated onto the robot platform. In the revise stage, it will assess whether the deduced solution was successfully executed or not, and add a success/fail flag to the case. In this way, the next time the solution is retrieved, the system will be able to exclude the toys that have failed as a play or a target toy. The retain stage maintains the play case-base by adding new cases and solutions or discarding those that are too similar to the ones that already exist. Maintaining cases and restricting the size of the case-base is important for real-time performance. Currently, the case-base size is restricted to 100 play cases.

IV. Results

A. *Experimental Setup*

To evaluate the proposed CBR turn-taking framework, we invited four subjects to manipulate given toys and video recorded them during the sessions. Two children, a 5-year-old and an 8-year-old, and two adults participated in the experiment. In each case, the subject was simply asked to perform stacking or inserting play with toys in the scene. A total of 36 video sessions were collected, and their lengths varied from 6 seconds to 75 seconds consisting of 3 to 9 turns, e.g., 5 stacking plays in a row.

Five of each stacking and inserting plays were randomly picked among the video sequences, and were used as a training phase. The training phase is a period dedicated to collecting cases to create an initial play case-base. The play case-base we have created and evaluated throughout this paper is a single database that consists of various stacking and inserting play cases.

When applied to play therapy, the training phase will have two meanings:

1) A play initiated by the robot:

When there is a clear therapy objective, a specific play case-base is inputted to the robot, and the robot guides the child to engage in a play by initiating a turn. Therefore, the training phase is placed before the session. In this approach, the robot has the potential to assess the child's play behavior and give feedback. (Fig.1. (a))

2) A play initiated by the child:

When the child is ready to lead his own play, the robot learns the child's play by building up the play case-base while observing. The training phase starts at the beginning of the session, and ends when the robot has collected a

TABLE I
TIME DISTRIBUTION FOR PBR AND TTCBR SYSTEM

(unit: ms)

PBR	ttCBR			Total
	Retrieve	Reuse	Revise /Retain	
~640	120~200	80~130	~0.20	840~970

sufficient amount of play cases. In this approach, the child is in charge of the play, and the robot adapts to the play as a passive playmate, imitating the child partner. (Fig.1. (b))

The scenes prepared for these experiments were occupied by 4 to 10 play objects on a table, and were not too overcrowded. Overlapping and occlusion of same colored toys were minimized at an initial setting unless relocated by subjects. The toys that were used in the experiment were blocks with a bucket, stacking rings, nesting cups, and a wooden block set.

B. Evaluation

Evaluating whether a deduced solution is viable or not can be quite tricky. The toys picked by the system can be different from those that were chosen by the human counterpart, but they might still be a good alternative. Therefore, instead of narrowing down to a single solution, we define a group that consists of solutions with sufficient similarity. If the actual choice made by the subject falls within the solution group, we evaluate it as successful. Grouping threshold and size are introduced to control the boundary that determines how strict we will be in identifying the toy as a match. For example, if grouping threshold is 0.2, toys with shape descriptors that are within 20% of the best matching descriptor are all considered as viable solutions.

Fig. 5 compares various play scenarios executed by the subject and the result of the ttCBR framework. The play scene is described in plain language for the reader's easy understanding, but each toys shape such as cup, block, and bin are registered in the system as shape descriptors as explained in Fig. 3. After the play scene has been converted into a play case, ttCBR retrieves the best matching case and its solution from the play case-base. During the reuse step, ttCBR produces group solutions, which are shown inside the brackets in Fig. 5 `Reuse`. While ttCBR successfully formed a group solution in the first two play scenarios, it missed a possible candidate during the third example. The grouping threshold determines the false positive and false negative rates. With the play case-base evaluated in this paper, we have found that setting grouping threshold to approximately 0.35 and limiting the group size to 3 gives the best average successful solution rate (82.36%) while keeping the false positive rate reasonably low.

The average run-time of our application for generating a solution is within a second, and has achieved real-time performance. Time distribution is depicted in Table I. Results show that our turn-taking Case-Based Reasoning framework delivers good strategy while maintaining robust real-time performance.

V. CONCLUSION

We have presented a methodology for planning turn-taking strategies using Case-Based Reasoning for a development of a therapeutic robot playmate. Our approach consists of building a play case structure that describes a play turn, retrieving a solution from similar past stored cases, and reusing it to adapt to the new play scene. The contribution of our work is in applying the Case-Based Reasoning concept to a child-robot turn-taking play, which enables the whole system to bypass a long complicated decision process. The proposed framework was able to deduce a solution within a second with a successful solution rate of 82.36%.

REFERENCES

[1] J. Piaget, *Play, dreams and imitation in childhood*. New York: Norton, 1962.

[2] S. Baron-Cohen, H. Ring, S. Wheelwright, E. Bullmore, M. Brammer, A. Simmons, and S. Williams, "Social intelligence in the normal and autistic brain: an fMRI study," *European Journal of Neuroscience*, vol. 11, no. 6, pp. 1891–1898, 1999.

[3] H. W. Park, A. Howard, and C. Kemp, "Understanding a child's play for robot interaction by sequencing play primitives using hidden markov models," in *Robotics and Automation, 2010. ICRA '10. IEEE International Conference on*, May 2010.

[4] G. Baranek, C. Barnett, E. Adams, N. Wolcott, L. Watson, and E. Crais, "Object play in infants with autism: Methodological issues in retrospective video analysis," *American Journal of Occupational Therapy*, vol. 59, no. 1, pp. 20–30, 2005.

[5] B. Robins, K. Dautenhahn, and P. Dickerson, "From isolation to communication: a case study evaluation of robot assisted play for children with autism with a minimally expressive humanoid robot," *Proc. Second Inter. Conf. Advances in CHI, ACHI'09.*, 2009.

[6] A. Billard, B. Robins, J. Nadel, and K. Dautenhahn, "Building robota, a mini-humanoid robot for the rehabilitation of children with autism," *RESNA Assistive Technology Journal*, vol. 19, no. 1, pp. 37–49, 2006.

[7] B. Scassellati, "How social robots will help us to diagnose, treat, and understand autism," in *Proc. of the 12th International Symposium of Robotics Research (ISSR05)*, 2005.

[8] A. Duquette, F. Michaud, and H. Mercier, "Exploring the use of a mobile robot as an imitation agent with children with low-functioning autism," *Autonomous Robots*, vol. 24, no. 2, pp. 147–157, 2008.

[9] P. Marti, A. Pollini, A. Rullo, and T. Shibata, "Engaging with artificial pets," in *EACE*. University of Athens, 2005, p. 106.

[10] A. Aamodt and E. Plaza, "Case-based reasoning," *Proc. MLnet Summer School on Machine Learning and Knowledge Acquisition*, pp. 1–58, 1994.

[11] T. Bushey, "Autism Games," http://sites.google.com/site/autismgames/home/games-pages/games-index-2-intermediate-level/your-turn, 2009, [Online; accessed 19-Jan-2010].

[12] K. Toth, J. Munson, A. N. Meltzoff, and G. Dawson, "Early predictors of communication development in young children with autism spectrum disorder: Joint attention, imitation, and toy play," *Journal of Autism and Developmental Disorders*, vol. 36, no. 8, pp. 993–1005, 2006.

[13] T. Charman, J. Swettenham, S. Baron-Cohen, A. Cox, G. Baird, and A. Drew, "Infants with autism: An investigation of empathy, pretend play, joint attention, and imitation," *Developmental psychology*, vol. 33, no. 5, pp. 781–789, 1997.

[14] G. Dawson, A. Meltzoff, J. Osterling, and J. Rinaldi, "Neuropsychological correlates of early symptoms of autism," *Child development*, vol. 69, no. 5, pp. 1276–1285, 1998.

[15] G. Dawson and A. Adams, "Imitation and social responsiveness in autistic children," *Journal of Abnormal Child Psychology*, vol. 12, no. 2, pp. 209–226, 1984.

[16] G. Dawson and L. Galpert, "Mothers' use of imitative play for facilitating social responsiveness and toy play in young autistic children," *Development and psychopathology*, vol. 2, no. 02, pp. 151–162, 2008.

[17] C. Breazeal, "Toward sociable robots," *Robotics and Autonomous Systems*, vol. 42, no. 3-4, pp. 167–175, 2003.

[18] C. Sidner, C. Lee, C. Kidd, N. Lesh, and C. Rich, "Explorations in engagement for humans and robots," *Artificial Intelligence*, vol. 166, no. 1-2, pp. 140–164, 2005.

[19] D. Lowe, "Distinctive image features from scale-invariant keypoints," *International journal of computer vision*, vol. 60, no. 2, pp. 91–110, 2004.

[20] H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," *Computer Vision–ECCV 2006*, pp. 404–417, 2006.

[21] R. Brunelli, *Template matching techniques in computer vision*. Wiley, 2008.