be within close proximity to the enemy. For ranged attacks this requirement did not exist, making attacks easier to complete. This explains the reduced damage values in Figure 3 for the warrior agent-only and control entries. For long ranged attacks, the ranger agent competed at a level comparable with human play, although it incurred more deaths on average. Finally, the wizard agent inflicted more damage than its human counterparts and also incurred fewer deaths. This is likely due to the behavior initially witnessed during trace capture, where human wizard players typically adopted a strategy which used the wizard's abilities to protect the group, while consistently avoiding the enemy. This behavior was successfully replicated by the behavior engine. The fact that hybrid human-agent teams outperform human-only teams at task completion is likely due to the prioritized case-base search, which ensured higher quality case-bases were queried first.

The authors were not expecting the hybrid human-agent groups to perform any better than the human-only groups in the qualitative evaluation. However, the results from Table 3 showed that for a number of questions the human-agent group members were less critical and more positive of their team compared to the human-only and control teams. In particular, hybrid human-agent groups performed the best on the final question: "*Generally speaking, I am very satisfied with this team*".

All of the subjects involved in the study thought that the member of the research team was making and executing playing decisions for their character, when in fact this character was controlled by the case-based behavior engine. After revealing to participants that they had been playing with an agent rather than a human team mate, a few subjects stated that they were so focused on their own actions they did not have the opportunity to fully observe the actions of the other players. One possible explanation for why human-agent groups were more satisfied with their team is that, given this high *cognitive load* subjects had to deal with, they may have relied more on the eventual outcome of the task to influence their responses to the questionnaire. The human-agent group had a 100% success rate at task completion, compared with lesser success rates for the human-only and control teams. This may have led to an improved qualitative judgment by team members.

## Conclusion

This work involved the design and construction of a case-based behavior engine, which learns from demonstration. Traces collected from multiple human demonstrations are used as input into the behavior engine, which then suggests a sequence of actions to perform based on those traces. By applying learning from demonstration, new behaviors can very quickly and easily be incorporated into the system without requiring additional programming effort. We utilized our behavior engine within a study to determine differences in performance between teams composed entirely of human players and teams augmented with artificial agents. We designed our study, such that human team members who played on the hybrid human-agent team, were not aware of this fact. This was done in order to not introduce any bias into the study. We found that hybrid, human-agent teams, were more successful than all-human teams at task completion. We also found that on some qualitative dimensions, hybrid teams were perceived and evaluated better by their human team members, compared to human-only teams.

## Acknowledgments

The authors wish to thank Danny Jugan and Axis Games for kindly providing the source code for Hands of War 2.

This material is based upon work supported by the National Science Foundation under Grant No. (IIS-1216253). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

## References

Aamodt, A., and Plaza, E. 1994. Case-Based Reasoning: Foundational issues, methodological variations, and system approaches. *AI Communications* 7(1):39–59.

Abraham, A. T., and McGee, K. 2010. AI for dynamic team-mate adaptation in games. In *Proceedings of the 2010 IEEE Conference on Computational Intelligence and Games, CIG 2010*, 419–426.

Axis-Games. 2012. Hands of war 2. http://armorgames.com/play/10987/hands-of-war-2.

Floyd, M. W.; Davoust, A.; and Esfandiari, B. 2008. Considerations for real-time spatially-aware case-based reasoning: A case study in robotic soccer imitation. In *Advances in Case-Based Reasoning, 9th European Conference, ECCBR 2008*, 195–209.

Merritt, T. R., and McGee, K. 2012. Protecting artificial team-mates: more seems like less. In *CHI Conference on Human Factors in Computing Systems, CHI '12*, 2793–2802.

Ontañón, S.; Mishra, K.; Sugandh, N.; and Ram, A. 2007. Case-based planning and execution for real-time strategy games. In *Case-Based Reasoning Research and Development, 7th International Conference on Case-Based Reasoning, ICCBR 2007*, 164–178.

Palma, R.; Sánchez-Ruiz-Granados, A. A.; Gómez-Martín, M. A.; Gómez-Martín, P. P.; and González-Calero, P. A. 2011. Combining expert knowledge and learning from demonstration in real-time strategy games. In *Case-Based Reasoning Research and Development - 19th International Conference on Case-Based Reasoning, ICCBR 2011*, 181–195.

Riesbeck, C. K., and Schank, R. C. 1989. *Inside Case-Based Reasoning*. Hillsdale, NJ, USA: L. Erlbaum Associates Inc.

Wageman, R.; Hackman, J. R.; and Lehman, E. V. 2005. Team Diagnostic Survey: Development of an instrument. *Journal of Applied Behavioral Science* 41(4):373–398.

Weber, B. G.; Mateas, M.; and Jhala, A. 2012. Learning from demonstration for goal-driven autonomy. In *Proceedings of the Twenty-Sixth Conference on Artificial Intelligence (AAAI-12)*.