Adam Goodkind
PhD Dissertation Summary

Predicting Social Dynamics in Interactions Using Keystroke Patterns

Every day, we communicate through computers on projects ranging from a group lunch order to booking a flight to learning critical medical information. And every day, we also *mis*communicate through computers: We don't pick up on an intentionally humorous response, or we miss the criticality of a request. This is made more frustrating because if these responses or requests were made in a face-to-face setting, these underlying intentions would be easier to pick up through tone of voice or the rate of speech, i.e. spoken prosody (Pierrehumbert and Hirschberg, 1990).

The COVID-19 pandemic, and its effects on remote working, have added a tragic emphasis to the need for a better understanding of computer-mediated communication, as text-based CMC has come to occupy an even more central role in our lives (Microsoft, 2021; Teevan et al., 2022).

My thesis aims to use timing patterns in typing, called *keystroke dynamics*, to detect underlying motivations and intentions, and make the information normally available only in face-to-face interactions also accessible in a text-based interaction, where prosodic information is assumed to be lost (Plank, 2016). If this information can be recovered and utilized, it will make text-based conversations more expressive and increase the bandwidth of information that it is possible to exchange via text.

For my thesis, I recruited 196 participants who took part in a 16-minute conversation where they exchanged movie and TV recommendations. I collected both the message texts as well as the timing information of every keystroke. Following the conversation, participants completed a survey that asked them to rate their opinion on aspects of their partner, as well as the conversation itself. My thesis uses all of these sources of information to investigate the underlying dynamics of a conversation.

The first study in my thesis used keystroke timing to infer characteristics of *dialogue acts* in a conversation, or the illocutionary function of an utterance. For example, some utterances, such as questions, are backward-facing dialogue acts that are intended to clarify a previous remark; on the other hand, forward-facing dialogue acts, such as a statement, presuppose that the existing conversation is agreed upon common ground and intend to advance the conversation (Stolcke et al., 1998). I use typing patterns to answer whether these different dialogue acts have different typing patterns associated

with them. This is important because a dialogue act such as a question would necessitate a very different type of reply compared to a statement. If a computer agent or a human interlocutor could gather more information about the type of dialogue act being produced, then they could also generate a more appropriate response.

Study 2 looks at adjacent pairs of utterances and infers underlying sentiment changes as well as the effect of a participant's opinion of their partner. A unique aspect of typing in dialogue, as opposed to in isolation, e.g. answering essay prompts or typing a thesis, is that a participant's utterances are dependent on, among other factors, previous context as well as the participant's overall impression of their partners. In my first experiment I investigate whether typing patterns can predict the sentiment of an utterance *per se*, as well as predict the change in sentiment from a preceding utterance to the current utterance. In the second experiment I look at whether the sentiment of a specific utterance as well as a participant's overall impression of their partner each exert an independent influence on typing patterns. Both of these experiments are important because during an interaction it is important to understand not only the words an interlocutor is producing, but also the underlying emotions an interlocutor is experiencing. This is the primary concern of the burgeoning field of *affective computing* (Buker and Vinciarelli, 2021; Picard, 2000): Not only is it important to understand the sentiment of a single utterance, but also how shifts in sentiment are manifested, as well as the overall emotions of a user.

Finally, my third study looks at the complex sentiment of rapport (Tickle-Degnen and Rosenthal, 1990), and how well a neural network can predict low rapport between partners. Because rapport is multidimensional, keystroke patterns provide an ideal production modality given that their patterns are sensitive to a number of influences, both social and cognitive. In a series of experiments, I test the predictive power of the full set of keystrokes, as well as subsets based on the participant's role in the conversation and subsets based on temporal slices. While the temporal subsets provide roughly the same amount of predictive accuracy, the subset of keystrokes collected when a participant is receiving recommendations is especially accurate. Predicting low rapport of a receiver is important in many settings, such as patient/provider or IT professional/user, where it is essential to maintain high rapport so that provided recommendations are well-received. Keystroke patterns allow for a continuous, non-obtrusive method for monitoring these feelings of rapport.

My findings have implications for bandwidth-mediated theories of computer-mediated com- munication, as well as channel expansion theories (Gergle, 2017; Walther, 2011, *inter alia*). In addition, my studies are based on the cognitive concept of Implicit Prosody (Fodor, 2002b), and so my findings could provide support for this theory. Finally, my

findings also bring up many ethical issues surrounding keystroke monitoring, and these are discussed as well.

Overall, my studies show how keystroke patterns are sensitive to a number of social dynamics and can detect signals that lexical information alone is less sensitive to. Uncovering these signals and sharing them with interlocutors, whether humans or computer agents, can improve the expressiveness of a conversation.

As a final note, my data as well as the code used to collect the data are publicly available at https://github.com/angoodkind/KiDcorpus. This corpus of data should be valuable to researchers in many different areas, and can be used to expand upon the foundations established in my thesis.