Digital-Microfluidic Biochips for Quantitative Analysis: Bridging the Gap between Microfluidics and Microbiology*

Mohamed Ibrahim and Krishnendu Chakrabarty

Department of Electrical and Computer Engineering
Duke University, Durham, NC 27708, USA
{mohamed.s.ibrahim, krishnendu.chakrabarty}@duke.edu

Abstract—Digital-microfluidics technology has shown considerable promise for advancing sample preparation and point-of-care diagnostics; therefore, it has the potential to transform microbiology and biochemistry research. Over the past decade, a number of microfluidics design-automation techniques have been developed for on-chip droplet manipulation. However, these methods overlook the myriad complexities of biomolecular protocols and they have yet to make a significant impact in biochemistry/microbiology research. A paradigm shift in biochip design automation and a “phase transition” in research are clearly needed to bridge this gap between microfluidics and microbiology. In this paper, we explain how researchers from design-automation and embedded systems can play a key role in this transition. We present a new synthesis flow that uses realistic models of biomolecular protocols and cyberphysical adaptation to address real-world microbiology applications. We also present a list of metrics that can be used for the assessment of design-automation techniques for microbiology applications.

I. INTRODUCTION

In recent years, digital microfluidics (DMF) has rapidly emerged as a key lab-on-a-chip technology that opens up new opportunities for point-of-care applications, including clinical diagnostics [1] and electrochemistry [2]. This technology has been recently deployed by Genmark Diagnostics for infectious disease testing [3] and FDA approval is expected soon for an analyzer from Baebies to detect lysosomal enzymes in newborns [4]. The primary advantages of this technology are automation, reduction of the consumed samples and reagents, and decrease in the reaction time due to the increase in surface-to-volume ratio [5]. These attributes (and others) are making DMF an increasingly popular platform for miniaturizing quantitative biomolecular applications ranging from blood glucose test [6], gene-expression analysis [7], [8], and pathogen detection [9]. Using digital-microfluidic biochips (DMFBs), bioassay reactions are scaled down to droplet size, and the control of nanoliter droplets for conducting fluid-handling operations (e.g., dispensing, transporting, mixing, and heating) is performed using electrowetting [10]. Reaction droplets are sandwiched between a grounded top plate and a bottom plate that contains an array of insulated driving electrodes, thus allowing bioassay protocols to be electrically reconfigurable on-the-fly.

This configuration has also enabled the seamless integration of sensors and imaging techniques and led to the first generation of cyberphysical DMFBs [11], [12], and software-based dynamic adaptation in response to sensor feedback has been utilized for error recovery [13]–[15]. Due to the fundamental importance of genomic analysis, considerable effort has similarly been devoted to the design and implementation of miniaturized platforms for gene-expression analysis [7], [16]–[18].

However, a drawback of previously reported platforms is that they were optimized for sample-limited analyses. They are inadequate for running downstream quantitative analysis, including the processing of multiple samples through independent pathways. Microfluidics design-automation (“synthesis”) techniques for on-chip droplet manipulation have been actively studied in recent years [19], but they overlook the myriad complexities of biomolecular protocols and they have yet to make an impact in biochemistry/microbiology research. It is therefore unlikely that today’s synthesis techniques will be able to cross the formidable barrier that separates design automation from practical biomolecular analyses.

The realization of the second-generation of cyberphysical DMFBs for quantitative biomolecular analysis requires a new synthesis flow that is based on the realistic modeling of biomolecular protocols. Our recent studies have led to the insight that the “sequencing graph” model [20], [21], which is based on a data-flow perspective of bioassays and forms the basis for today’s DMFB synthesis solutions, is insufficient for realistic biomolecular protocols. This shortcoming is highlighted by the fact that experimental demonstrations of run-time adaptation in our work on first-generation cyberphysical DMFBs did not consider control flow to incorporate the specific “if-then-else” requirements of multiple sample pathways [22], [23]. In other words, a capability is needed to determine which pathway must be followed to achieve successful completion of a bioassay based upon the monitored status of the assay during run time.

In addition, preliminary studies also highlight the need to empower cyberphysical DMFBs with an interactive firmware.
layer that can collect data from on-chip sensors, perform real-time data analysis, and guide resource allocation with appropriate decisions [24]. However, the transition to such an autonomous platform is impeded by the lack of design support for intelligent real-time decision-making in a multi-assay setting. Techniques developed thus far are static; runtime adaptation based on assay outcomes is not supported and only a rigid path of bioassays can be run on the chip. There is a need to incorporate decision-making based on prescribed criteria; this breakthrough will ensure that a diverse collection of protocol paths can be traversed. Advances in design techniques for these platforms will push the frontiers in several application areas, including genome-wide screening, epigenetic inheritance and cancer research, and proteomic analysis.

This paper is therefore motivated by the need to enable bench-top biomolecular assays on programmable and cyber-physical DMFBs. We explain how researchers from design-automation and embedded systems can play a key role in this transition. We present a new synthesis flow that uses realistic models of biomolecular protocols and cyberphysical adaptation to address real-world microbiology applications. The discussion in the paper is centered on realistic case studies involving quantitative gene expression analysis and epigenetics. We present a proposal on a standardized list of metrics that can be used for the assessment of design-automation techniques. We believe that this work will serve as a “call to arms” for more focused and relevant research to increase the adoption of digital microfluidics in translational research for point-of-care diagnostics, as well as for the detection and treatment of diseases such as cancer.

The rest of the paper is organized as follows. Section II describes characteristics of quantitative protocols and their modeling technique. Next, we introduce the new design-automation framework and describes the use of real-time system (RTS) control in Section III. A brief discussion on the assessment of design-automation methods is presented in Section IV and conclusions are drawn in Section V.

II. CHARACTERISTICS OF QUANTITATIVE BIOMOLECULAR PROTOCOLS

Quantitative-analysis protocols rely on the estimation of the dose-response relation curve for biomolecular compounds (e.g., protein, gene, cell, tissue, and organ), wherein the response of a biomolecular compound is studied with changing dose. Using the dose-response curve, we can estimate the dose or concentration of a substance, such as dye fluorescence intensity, associated with a specific biological response, such as gene expression. Fig. 1 shows an example of a dose-response relation curve for investigating blood-glucose concentration using a colorimetric assay: the change of dye color is associated with glucose concentration and it is detected using an on-chip absorbance measurement system such as LED-photodiode [6]. In this context, the dose-response relation curve is known as glucose-calibration curve. This curve therefore helps in interpolating the concentration of a glucose sample under test. As shown in Fig. 1, the reaction rate of the sample is a point on the Y-axis and the corresponding point on the X-axis is the sample concentration. Similarly, the dose-response relation is also applied through quantitative polymerase chain reaction (qPCR) to investigate gene expression, producing what is known as qPCR standard curve [25].

Our goal is to develop a design-automation framework to execute this class of protocols using autonomous cyberphysical DMFBs. For this purpose, it is necessary to pinpoint the main characteristics of these protocols, whereby these characteristics can be incorporated into the algorithms underlying the design-automation flow. Here, we list the main characteristics and the associated modeling schemes.

A. Independent Sample Pathways & Quantification Results

Quantitative results are reported numerically and are compared against an accompanying reference interval for interpretation. For this purpose, multiple samples are treated through independent sample pathways. Subsequently, the distributed detection results of the sample pathways are collected and fused to report a final quantitative result. For example, to study the transcriptional profile of a green fluorescent protein (GFP) reporter gene under epigenetic control, gene-expression analysis is conducted using three types of S. pombe samples: (i) control (GFP not under epigenetic control) samples; (ii) experimental (GFP under epigenetic control) samples; (iii) wild-type samples used as a reference to improve the outcome efficiency. Note that these samples must be run independently in a set of reactions in which the investigated gene is GFP [26].

To support multiple sample pathways, we extend the traditional sequencing graph model to represent the associated bioassay protocols for all samples, as shown in Fig. 2(a). Note that these protocol procedures can either be identical for all samples [8] or be unique to each sample type according to the application specifications. This modeling technique is similar to the multifunctional design scheme presented in [27],
which combines a set of multiplexed bioassay protocols in the same platform. Our solution, however, executes the bioassay protocols concurrently and collects detection results from all these protocols (using firmware) to reach a meaningful conclusion.

B. Uncertainty about Pathway Sequence

A second important feature of quantitative protocols, especially epigenetic protocols, is the inherent uncertainty about the order of basic fluidic steps for each sample. In other words, our recent benchtop experience with multiple sample pathways highlights the need for incorporating decision-making and adaptation capability into our framework [25]—such “if-then-else” reconfiguration can be performed either automatically using cyberphysical integration or using human-in-the-loop if necessary.

As shown in Fig. 2(b), the previous protocol model is modified to account for decision-making situations; a bioassay protocol for each sample pathway is decomposed into a set of supernodes, and an edge linking two supernodes represents a potential transition (decision-making) path. A supernode may represent either an entire bioassay or a part of a bioassay after which a decision must be made. For example, nucleic-acid (NA) isolation during the execution of qPCR-based gene-expression analysis must be followed with a decision on whether the quality of the isolated NA allows subsequent cDNA synthesis or requires further steps to get rid of the residual protein and debris.

C. Temporal Constraints

Temporal constraints may arise during protocol execution due to physical phenomena such as droplet evaporation or deadlines imposed by the target chemistry, e.g., degradation of samples and reagents [28]. Therefore, we need to provide a systematic approach that handles such constraints during the course of an experiment in order to ensure robust and reproducible quantitative results. An example of an upper-bound temporal constraint can be seen in chromatin immunoprecipitation (ChIP) protocol, in which the protein is initially cross-linked to DNA. It has been shown that an excessively long cross-linking time results in the majority of chromatin being resistant to shearing, thus degrading chromatin shearing process [29]. Therefore, an upper-bound temporal constraint on cross-linking reactions must be specified; this temporal constraint is often cell type-dependent.

As shown in Fig. 2(c), temporal constraints are incorporated into the protocol model and formulated as a Simple Temporal Problem (STP) [30]. A dotted edge in Fig. 2(c) represents a temporal constraint, bounding the time difference between the two connected operations. Each of these dotted edges is labeled by a single interval $[L, U]$ to specify the lower and upper bounds. An unspecified constraint is equivalent to a constraint with an infinite weight; for example $[-\infty, U]$ indicates an unspecified lower-bound constraint. Several algorithms can be used to check the consistency of a given STP [30].

D. Avoidance of Sample Interference

The samples in an experiment exhibit different biological properties that need to be analyzed and distinguished. Obtaining precise results, however, requires processing these samples without causing biological interference (i.e., cross-contamination), especially in experiments involving reagents that may lead to biofouling. As a result, conducting such experiments using a resource-limited setting such as a DMFB needs stringent quality-control procedures to eliminate causes of cross-contamination. A potential approach is to carry out washing operations on demand and synchronize functional biochemical procedures with washing steps during the course of an experiment.

III. SYNTHESIS FLOW: TRANSITION TOWARDS A REAL-TIME SYSTEM APPROACH

Based on the characteristics described above, the realization of quantitative protocols using a resource-limited DMFB requires a synthesis flow that takes into consideration the dynamics of fluid-handling operations within multiple sample pathways. Conventional synthesis methods heavily rely on offline optimization techniques [19], thus they cannot be scaled or easily adopted to support real-time interactions. As a result, in order to fulfill the unique characteristic of such protocols, we need to develop an alternative synthesis approach.

The synthesis problem based on the above characteristics can be mapped into a real-time coordination problem between multiple samples and a set of heterogeneous devices. However, an efficient transition towards an RTS-based synthesis
A. System Model

In [31], Ibrahim et al. presented the first work to model a digital-microfluidic system for quantitative epigenetics (e.g., gene-regulation analysis) in terms of real-time computing systems. Using the protocol model in Fig. 2(c), a supernode (i.e., a bioassay) is represented as a fluidic task \( \tau_b \), and the constituent fluid-handling operations are mapped into a set of fluidic substasks \( \{ \tau_{b1}, \tau_{b2}, ..., \tau_{bm} \} \). The formal definition of the fluidic task set \( T = \{ \tau_1, \tau_2, ..., \tau_n \} \) is also considered in [31]; this model captures the temporal and interdependency relationships among chemistries. The interaction between the task set \( T \) and a set of heterogeneous resources is also formulated as a mixed-integer linear program (MILP), enabling design-automation researchers to have an optimal reference to compare their solutions with. An optimal solution achieves minimum completion time while satisfying problem constraints.

B. MILP-Based Synthesis

A conventional approach to address the synthesis problem is to solve the MILP using a software optimizer [32], leading to optimal solutions for operation assignment, scheduling, and droplet routing (Fig. 3(b)). Despite optimality, the complexity of this approach is formidable and computational intractability is inevitable for practical quantitative protocols. Therefore, the adaptation of protocol flow based on sample-dependent decision making is not feasible if MILP-based synthesis is employed.

C. Hybrid Synthesis

In [33], Gombolay et al. developed a fast scheduling framework for a multi-robot setting that involves spatio-temporal constraints. In this framework, the coordination problem, formulated as a MILP, is solved by decomposing the MILP into a task allocation and a task scheduling subproblem. A hybrid solution is then generated by solving a simplified MILP for task assignment in conjunction with a dynamic real-time scheduler. This approach can generate near-optimal task assignment and schedules, and it is applicable to dynamic and large-scale problems.

Similarly, synthesis of DMFBs for quantitative analysis can be solved by decomposing the MILP into three subproblems: fluidic-operation assignment (\( FA \)), fluidic-operation scheduling (\( FS \)), and physical-level synthesis (\( PS \)); see the boundary lines in Fig. 3(c). A synthesis solution is then obtained using a hybrid approach such that the MILP is simplified and highly efficient results are obtained in a reasonable time. This approach, however, entails two main challenges:

1. The MILP formulation leads to strong coupling between the subproblems; thus decomposition requires a thorough investigation to ensure that the hybrid solution converges to the optimal case. In other words, performance-responsiveness trade-offs need to be studied.

2. A variety of hybrid solutions can be generated as each subproblem can be solved either optimally or heuristically. For example, a possible method uses MILP solver to solve the simplified MILPs for \( FS \) and \( PS \), whereas another method solves only \( FA \) optimally.

Despite these challenges, hybrid synthesis is a promising approach that is likely to gain more attention from design-automation researchers.

D. Dynamic Real-Time Synthesis

Obtaining a solution for MILP instances is often computationally expensive even after simplification. Therefore, an alternative approach for synthesis is to leverage real-time multiprocessor scheduling; such a technique has been proposed in [31]. This approach, despite being sub-optimal, scales efficiently to multiple independent biological samples.
and supports on-the-fly adaptation under dependency and spatio-temporal constraints, using a dynamic heuristic method (Fig. 3(d)). A simple dynamic policy is adopted allowing chemistries to be monitored and prioritized based on their progress towards deadlines.

E. Hierarchical Real-Time Microfluidic Systems

In [31], Ibrahim et al. proposed hierarchical system scheduling to support the timely coordination between synthesis execution, the periodic loading of actuation sequences, and firmware computation. This framework can be advanced by incorporating a component that monitors biochip utilization and invokes a washing procedure whenever needed to counter sample interference (discussed in Section II-D). A finite-state machine (FSM) is designed to keep track of the status of each electrode (see Fig. 4(b)); such a mechanism relies on information communicated from the synthesis component during runtime and it aims to synchronize washing operations with functional reactions to minimize washing overhead. The hierarchy of components for the modified real-time microfluidic system is illustrated in Fig. 4(a). Note that unlike the washing method in [34], our washing mechanism is more flexible as it is sample type-dependent.

IV. ASSESSMENT OF SYNTHESIS TECHNIQUES

In this section, we introduce a methodology for evaluating the synthesis techniques for implementing quantitative biomolecular protocols. We examine representative benchmarks and a set of metrics for comparison.

A. Benchmarks

The majority of complex biomolecular protocols are constructed of well-defined building blocks such as qPCR. For instance, gene-regulation analysis is a complicated protocol that uses stages of qPCR-based gene-expression analysis and ChIP [31]. Therefore, design-automation researchers can streamline their assessment methodology by using these blocks or a combination of them. Graphical representation of gene-expression analysis and ChIP can be found in [8], [35].

B. Metrics

Synthesis frameworks can be assessed based on four metrics: completion time, adaptation time, protocol attributes, and biochip lifetime.

1) Completion time: This metric refers to the total time elapsed from the beginning to the end of a benchmark protocol. Completion time is measured in terms of the system actuation clock period, and it assumes a pre-specified set of sample pathways. This metric can be utilized to evaluate the performance of a proposed synthesis mechanism and its ability to efficiently perform real-time scheduling.

2) Adaptation time: This metric refers to the computation time necessary to adapt a protocol flow, i.e., to generate a new synthesis solution on-the-fly, due to a decision taken during the analysis of a sample pathway. Adaptation time is measured in terms of milliseconds, and it can be used to assess the responsiveness and scalability of a synthesis framework—scalable frameworks tend to be responsive to protocol changes (i.e., exhibit short adaptation time), regardless of the number of sample pathways and the complexity of the protocol model.

As shown in Fig. 5(b), the dynamic real-time synthesis method is highly responsive since it relies on a heuristic approach. The MILP-based method, on the other hand, is computationally expensive and it requires the longest adaptation time; thus it might be impractical for large-scale biomolecular protocols.

3) Protocol attributes: We use this metric to evaluate whether a synthesis framework captures certain attributes of protocol chemistry, such as biological interference and reagent degradation. For example, the dynamic real-time synthesis in [31] prevents reagent degradation by including pathway-related temporal constraints and replenishment operations. However, biological interference has not been considered in this work. This metric is therefore used to investigate the efficiency of system modeling in a synthesis framework, rather than the algorithmic innovation.
4) Biochip lifetime: Excessive usage of the specific electrodes that control on-chip resources such as heaters or detectors may lead to degradation, which significantly impacts biochip lifetime [25]. As a result, synthesis techniques can be designed to incorporate technology-specific degradation models [18], thus enhancing the biochip’s lifetime and reducing the likelihood of system failure. The work in [8] uses spatial reconfiguration to reduce biochip degradation during synthesis, which is measured in terms of the maximum occupancy time of an electrode. The degradation-aware synthesis in this work, however, considers the degradation constraint as a soft constraint. Hence, it does not guarantee that resources with completely degraded (i.e., broken) electrodes will be avoided; as a result, failures are possible during protocol execution.

V. CONCLUSION

In this paper, we have introduced a new synthesis methodology for DMFBs to bridge the existing gap between microfluidics and biomolecular protocols. The transition to the proposed synthesis methodology has been motivated by the identification of important aspects of biomolecular protocols, and various solutions have been discussed to fulfill domain-specific aspects. Finally, we have presented an assessment methodology that motivates design-automation researchers to evaluate their work and compare between various solutions. This paper hence represents a key step forward towards advanced microfluidics technology for realistic on-chip biochemistry.

REFERENCES


